The Dissertation Committee for Emre Yucel
certifies that this is the approved version of the following dissertation:

Capacity Auctions for Electricity

Committee:

James Dyer, Supervisor

John C. Butler

Kumar Muthuraman

Edward Anderson

Capacity Auctions for Electricity

by

Emre Yucel

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

December 2018

Acknowledgments

I wish to acknowledge the guidance and support of my committee: Jim Dyer, John Butler, Kumar Muthuraman, and Ed Anderson. I would not be writing this without all of their help throughout the entire process. I especially want to thank Dr. Dyer and Dr. Butler, whose endless patience and persistence with me on this seven and a half year odyssey has been without a doubt the biggest driver of my success. I am extremely appreciative of all their efforts in pushing me to be a better thinker, researcher, writer, presenter, and storyteller.

I met so many amazing people: Suho Han, Shannon Chen, Ben Van Landuyt, Kathleen Powers, Prasart Jongjaroenkamol, Katerina Zukis, Zack Liu, and Will Dahlmeier. I am grateful for these friendships and all of the support, joy, and memories they have brought me. Especially Will Dahlmeier, who was a constant companion, a sounding board and springboard for so many of my ideas, and who continues to be a fantastic source of laughter and discourse.

I am appreciative of my coworkers and the entire team at HomeAway for their flexibility and accommodation as I finished my final year in the program.

Lastly, I would like to thank my parents Adnan and Mine and my sister Elif for all of their encouragement and support, not only during this journey, but in all of my pursuits.

Capacity Auctions for Electricity

Publication No. ____________________

Emre Yucel, Ph.D.
The University of Texas at Austin, 2018

Supervisor: James Dyer

Faced with uncertainty of future electricity generation supply, many regional electricity markets have adopted or considered adopting capacity markets for electricity. We study the structure of these markets and in particular capacity supply auctions such as the one implemented by PJM Interconnection (PJM), a regional transmission organization. Participants bid generation capacity into the auction, and those that win receive a capacity payment in return for having this capacity available for generation at a future delivery date. The auctions can be classified as multi-unit uniform price auctions, though price is set according to a demand curve rather than by participants' bids. We find closed-form solutions for the optimal bids as a function of cost, study welfare impacts of the auction, and show how the results can be extended numerically for more complex situations. We then use these optimal bid functions in an agent-based simulation of electricity markets, comparing energy-only markets to capacity markets and measuring the impact on both the generators and consumers of electricity. Lastly we use our agent-based simulation model coupled with reinforcement learners to determine whether or not the optimal bid strategy discovered in

the beginning can be learned over time by agents participating in the energy and capacity markets.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Faced with uncertainty of future electricity generation supply, many regional electricity markets have adopted or considered adopting capacity markets for electricity. The standard wholesale electricity markets widely used around the world consist of a real-time and day-ahead spot market that allows supply and demand to schedule load and meet deficiencies. Unfortunately, there have been situations in which these energy-only markets have been unable to efficiently clear for various reasons when demand is very high and further generating capacity is unavailable. Some of these include the introduction of artificial price caps by regulators, insufficient quantities of real-time demand response, and various other actions taken by the independent system operator (ISOs) that would be considered out of market (Joskow (2008)).

These high demand situations only happen for a few hours a year (usually in the hottest summer months, or coldest winter months), but generation capacity must be built to be able to supply electricity during these times. Unfortunately, this capacity is only earning money for a small portion of the available hours in a year, and could even go a year or more without being called upon to generate. For a power plant to recover their fixed costs, the price must be very high during the few hours that they operate. Base load generation has higher capital costs and lower operating costs, while peaking units tend to have lower capital costs but higher operating costs. The base load generation may make a large amount of its revenues by operating during the peak demand times; the peakers, however, must earn nearly all their revenue during this time. If the price is too low, because of price caps or otherwise, the revenues will be too low for these peaking units, and generation investment will fall. This of course will lead to inadequate capacity to deal with high demand, and in turn increase the probability of a loss of load event.

There have been a few suggestions to help alleviate the inefficiencies in the energy only market such as those presented by Joskow (2007). The price caps that are currently in place in most markets are lower than what the clearing price would be in a competitive market. Raising these price caps would allow the peakers to recover their costs when called upon. Whenever out of market actions are taken, the price should rise to the price cap. Tools like rolling blackouts or voltage reductions are used by the system operator but are not correctly reflected in the real-time market. Joskow notes that care must be taken to ensure that market power is mitigated. Those with market power can artificially induce scarcity, forcing the price to the price cap and receiving considerable sums of money.

In addition to reforming the wholesale energy markets, forward capacity markets have been proposed as a way to eliminate this missing money gap. In markets where there is a price cap or other barriers to prevent price reaching the optimum price during peak times (this would be a price that allows peaking units to recover their costs), then a capacity payment should be provided to cover the difference. Implementing this in practice is a different challenge.

In the early to mid 2000s, capacity payments mechanisms were implemented based on a target reserve margin which was determined by the ratio of generating capacity over peak demand. Unfortunately, the capacity payments calculated were too low to make up for the gap in revenue from participating in the wholesale electricity market. Additionally, based on a target reliability estimate and thus a target generation capacity, the price paid was zero when capacity was over the target generation, and a fixed, flat payment when below the target. This led to volatility in the capacity payment from year to year. Many system operators introduced a demand curve which made the capacity price a smooth decreasing function of target generation capacity. This type of demand curve style auction is what we will examine in this paper. In particular, we focus on the type of auction run by PJM Interconnection (PJM), which can be described as a sealed-bid, single-price, multi-unit supply auction.

The basics of the PJM auction are as follows. Each year an auction is held for electricity capacity delivery three years in the future (that is, the auction in 2015 will be for the 2018 delivery year). The ISO sets certain auction parameters, the two most important being the total peak forecast demand and the reliability requirement. The reliability requirement ensures that there is some excess percentage of capacity during peak demand. These two numbers form the basis for the quantities of PJMs demand curve for capacity. Prices are based on the cost of new entry, and set to represent the payment a new combustion turbine would need yearly to break even. PJM publishes a demand curve each year for the auction.

On the supply side, bids are submitted by generators indicating the capacity payment they would receive in exchange for the generation capacity provided. These bids are aggregated and a supply curve is formed. The intersection of this supply curve and PJMs demand curve determine the capacity payment received by those clearing in the market. There are more specifics, such as price differentials between regions, and further auxiliary auctions, but this simplified representation is all that will be necessary for framing the context of the problem.

The literature in standard auction theory is well-established. The seminal work by Vickrey (1961) showed that in a standard first price auction, participants have an incentive to shade their bids, while a second price auction is incentive compatible. Engelbrecht-Wiggans and Kahn (1998) examine the multi-unit uniform price auction. They find that when there are M units of goods for sale, each bidder wants to purchase two units, and not all bidders can be satisfied, bidders underbid for the second good and thus pay very low prices. Gretschko et al. (2014) prove that any equilibrium in a strictly descending multi-unit auction is inefficient. Swinkels (2001) shows that in discriminatory and uniform price auctions for multiple units, as the number of participants grows very large, the inefficiency in these auctions goes to zero. These auctions are the more common single seller, multiple buyer auctions, whereas we will discuss a single buyer, multiple seller auction.

The single buyer, multiple seller auctions presented in the rest of this paper resemble the three common types of multi-unit auctions (discriminatory, uniform-price, and Vickrey) as described by Krishna (2009); they differ in that the price is not determined entirely by bidders' bids, but rather it also depends on an exogenous demand curve specified by the auction holder. This acts as a reserve price that is unknown to the bidders. Ausubel and Cramton (2004) generalize the multi-unit Vickrey auction to allow for reserve pricing. They find that truthful bidding is a dominant strategy in this scenario. Failure to set a reserve price correctly, however, can lead to poor outcomes as described in Klemperer (2002). Blume et al. (2009) observe that for any positive reserve price bidders bid truthfully when their valuation exceeds the reserve price in the buyers auction. In a seller's auction this would mean that suppliers would bid truthfully when their cost is higher than the minimum allowable capacity price. Other work has been done to determine the optimal level of the reserve price, such as that by Larsen et al. (2004) and Cramton and Stoft (2007).

Both Hortasu and Puller (2008) and Schwenen (2015) analyze a multi-unit supplier auction such as the ones considered here. However, both assume perfect information on the part of the participants, both about payoffs and costs of other generators. Contrasted with the standard auction participant who receives their perceived value minus their cost, the auction participant in our paper receives some unknown payment that is a function of all bids minus their cost. This added uncertainty complicates the problem and has not been analyzed fully but has a large impact on bidding behavior as we will show. In their paper, Hobbs et al. (2007) show that the decreasing price auction results in lower costs to consumers and more stability in long term capacity prices than the flat price auction.

In Chapter 2, we analyze the capacity auction described above as a game in the economic sense. We begin with a simple example and work our way through progressively more complicated auction structures ending with the single buyer, multiple seller uniform-price auction. By the end of the chapter we will have analyzed a mechanism that is quite close in theory to that used by independent system operators in practice. We begin by first introducing the game of chicken in the context of a capacity auction to demonstrate the core concepts of bidding above and below cost, as well as the impact of multiple bidders. From there we begin the brunt of the analysis. In the second section we analyze a single generator in isolation in three scenarios. The first scenario is that when the clearing price is known, the second when it is unknown, and the third scenario when it is unknown and the game is repeated over multiple time periods. We find that in all cases truthful bidding is the optimal strategy, and furthermore, in multi-stage games, bidding 0 for every period after the first is optimal.

In the third section we introduce other bidders. At first, we consider a situation where the price is unknown and a single winner is chosen. The optimal bid is calculated in closed form as a function of the number of participants and a generators true cost. From there, the restrictions are lifted and multiple winners are permitted. This time however, a closed form solution is unable to be found, and a numerical solution is presented instead. The results are presented as a function of the number of winners holding total participants fixed, and as a function of the number of participants while holding the number of winners fixed. We find that when more bidders are present, the degree of bid-shading increases, and conversely, when the number of bidders is fixed and the number of winners is increased, the degree of bid-shading decreases.

We conclude the optimal bid functions section with a game where the price is not fixed, but is determined by the intersection of a demand curve set by the ISO and the bidders bids. We find a closed form solution for the optimal bid as a function of the number of participants and the slope of the demand curve. Here we find that the severity of bid-shading is less than before but still dependent on the number of participants and the slope of the demand curve.

In Section 4, the bid functions found in Section 3 are used to analyze the health and welfare of the market. The probabilities that the ISO receives new capacity at any given price is compared across the different scenarios described previously. Analogously, the probability that any individual generator clears as a function of clearing price is also presented. To compare the differing mechanisms, both the demand curve auction and a fixed price with fixed number of winners auction are analyzed. A reasonable scenario in which an ISO needs new generation capacity serves as the basis for calculating total costs to the system. Both capacity price payouts and costs as a result of lost load are measured and contrasted.

In Chapter 3 we aim to take the analysis performed in Chapter 2 using the theoretical model and apply it to a simulation model of electricity markets to see whether the theory holds at scale. To do so, we turn to a simulation model where each generator is modeled by a single agent. These agents can then be given a set of behaviors to follow and set off to interact with each other in a simulation of an electricity market. Researchers at Iowa State University developed the Agent-based Model of Electricity Systems (AMES) to perform such analyses on a day-ahead electricity market. Additionally, each generator in the simulation is equipped with a reinforcement learner module, allowing it to optimize its bids in this day-ahead market. Sun and Tesfatsion (2007) first use the model to test the performance of the wholesale power market platform (WPMP) proposed by the Federal Energy Regulatory Commission (FERC) in 2003. They implement a simple 5-node model to showcase the modeling framework and to examine the ability of generators to exercise market power under the WPMP rules. They find that when allowed to bid strategically, the generators do submit higher than true marginal costs to the ISO, and that overall costs are three times higher than if they had reported truthfully. However, there is lower volatility during peak periods and reduced congestion on the transmission lines.

The model developed by Sun and Tesfatsion does not allow for competitive market entry nor does it have any form of capacity market, only the day-ahead market. Therefore, it is not ideal for investigating the long-term impacts of or abilities for electricity markets to adapt to changing demand conditions and increasing load growth over time. We seek to add these components as modules to the existing AMES day-ahead market and learning framework.

The competitive market-based entry uses a levelized avoided cost of electricity (LACE) vs. levelized cost of electricity (LCOE) comparison to determine market entry criterion. This approach has been modeled in the National Energy Modeling System (NEMS) model of the Energy Information Administration (EIA) (Namovicz (2013)) and has been used in the Full Cost of Electricity report from the Energy Institute (Mann et al. (2017)). Every generation source has a LCOE associated with producing a MWh of electricity. Similarly, each source also has an average revenue LACE from producing a MWh of electricity. If the LACE is greater than the LCOE, that is to say the revenue is greater than the costs, the generator will choose to enter the market.

The capacity market uses a demand function of the kind implemented by PJM in their own capacity market. This type of capacity market has a target reserve margin set by the ISO. The reserve margin represents the amount of capacity held in reserve in case of outages. Each year PJM projects a peak demand number, and the total amount of capacity installed in the system divided by the peak demand is the reserve ratio. Subtracting 1 from this gives the reserve margin. Most ISOs aim to have a reserve margin in the range of 10-15%.

Capacity is paid a higher price when the installed reserve margin is low relative to the target and is paid less when the installed reserve margin is high relative to the target. The capacity auction is held yearly, and both existing and new generation bid in to the market. The price paid is the intersection of the demand curve and the aggregated capacity supply offers from all generators. Those that clear in the auction are paid this price per MW of capacity they submit into the market that year. There are penalties for non-performance, but we do not address such details here.

Because our analysis focuses on the impact of a capacity market on energy markets, we must give our agents a set of rules to follow when bidding into the auction. Generators bid in according to the rules determined in Chapter 2. Recall that this means truthfully for new generation and 0 for existing generation. We begin with the case of an energy-only market. The only way new generators can enter the market is when the necessary economic conditions (LACE > LCOE) are met. A very loose approximation to the Texas energy market is constructed, and the simulation run for a period of 10 years. The performance of the market is observed, from average electricity prices and generator revenues to installed reserve margins and loss of load events. The simulation is then repeated at various levels of maximum price caps.

Following the energy-only case analysis, a market with a capacity auction is tested. The results of this are compared with those of energy-only markets. We find that although capacity auctions are costlier at providing similar levels of electricity, the reliability that they provide can offset this cost at small relative cost to the consumer.

In Chapter 4 we return to the learner mentioned earlier. Our goal for this chapter is to examine whether or not the results found in Chapter 2 can arise naturally using a model for human behavior and the model of electricity markets presented in Chapter 3. Rather than taking the results from Chapter 2 as given, we allow generators to bid from a pool of actions (ranging from -100% to +100% of their actual cost). We use a reinforcement learner, which means that actions that result in favorable outcomes are more likely to be played in the future. Conversely, those that lead to unfavorable outcomes are less likely to be played. Each learner starts out with a prior distribution over the actions, and as the game is iterated, updates the actions according to some reward function that maps rewards to changes in the probabilities. We develop a reward function to appropriately model the returns of clearing or not clearing in the capacity auction. Each year in our model represents an iteration in the game. We hope to find that the probability distribution over the actions at the end of our simulation correspond closely to the behaviors we defined in Chapter 2.

Roth and Erev (1995) first employ this reinforcement learner to study repeated games. Their implementation of this learner is influenced by aspects of learning discussed in the psychology literature. They aim to show that a simple model of behavior can closely match the observed behavior of humans. The simulation model of the learner is compared with that of experimental data and shown to closely approach the observed behavior from the experiments. In a later paper (Erev and Roth (2007)) they again revisit the ability of agent-based learners to accurately model human behavior. They conclude that even simple agent-based learners such as the one implemented in this paper can capture numerous aspects of human behavior. It is with this endorsement that we feel confident in using the reinforcement learner to model generator bidding behavior in capacity auctions.

We begin the chapter by introducing the Roth-Erev learning algorithm and the assumptions made to support it. Afterward, our variation on the algorithm is introduced. The reward function and its multiple sources of signal updates are described in detail. Additionally, a simple example is provided so that the reader may follow along and fully comprehend the mechanism by which the signal updates occur. Following this exposition, the simulations are run. First, we permit only new generators to learn strategies, while existing generators must bid 0 as before. Subsequently, we then allow both existing and new generation to learn the optimal bidding behavior. The results of this Chapter show that the optimal strategies found in Chapter 2 and used in Chapter 3 hold true. The structure of the auction lends itself to truthful bidding as an optimal strategy for new entrants, and 0-bidding holds for existing entrants. We conclude with some final words and recommendations based on the research completed, as well as future directions and extensions that would be possible.

Chapter 2

Capacity Auctions as a Mechanism

2.1 A Simple Game

To motivate our development we consider two players in a market. These can be interpreted as the marginal bidders in the full supply auction. Both are offering in the same amount of generation capacity, and both have the same costs. For now, we ignore any price submitted with their bids, assuming that both will be accepted if they enter. If a generator does not enter, then obviously it will not receive a capacity payment and its total payoff is 0. If one enters and the other does not, the entrant's payoff is the capacity payment minus their cost, which we denote with h. If both enter, both receive the net amount $l < h$; the price is lower than when only a single generator enters, as more capacity is supplied to the market, moving further down the demand curve. The game can be seen in Figure 2.1. If l is positive, then the dominant strategy is to enter no matter what. The interesting case is when l is negative. This could be the case that the capacity payments minus their costs results in a financial loss. Then this becomes the well known game of chicken described by Rapoport and Chammah (1966).

	Enter	Don't Enter
Enter	l, l	$h, 0$
Don't Enter	$0, h$	$0, 0$

Figure 2.1: A simple game for two firms

This game has two pure strategies and one mixed strategy equilibrium. The two pure strategy equilibria are enter/don't enter for firm 1 and vice-versa for firm 2. The mixed strategy for either entrant has the probability of entering equal to

$$\frac{h}{h-l}, \quad \text{for } h > 0 > l. \tag{2.1}$$

The expected capacity additions would then be

$$\frac{2sh}{h-l}, \quad \text{for } h > 0 > l, \tag{2.2}$$

where s is the amount of generation capacity being entered into the market per participant. For a market operator to ensure acquisition of new capacity it would be best to set market parameters such that h is high and l is as small as possible. This can lead to high capacity payments, which in the end are passed on to the consumer. There is a balance that needs to be struck between total capacity payments and capacity acquired.

2.2 Preliminaries

In the previous section, we assumed the bidders had identical costs, and did not account for the bidding in of costs that occur in actual capacity auctions. In this section we will define frequently used terms and the basic structure of the auctions. Henceforth, bidders will bid w in dollars per megawatt that represents their cost of producing capacity to generate electricity. The capacity price p in \$/megawatt is paid to bidders who win the auction for the amount of capacity they bid into the auction. We will commonly refer to capacity price and capacity payment as the same thing throughout this paper. Bidders also submit a quantity q of capacity supply along with their bid w, but in this paper we will not be focusing on the quantity portion of bids, choosing to model bidders of identical sizes.

The ISO has a demand curve for capacity, with which they determine the price for capacity. This acts in essence like a reserve price: if the supply of capacity offered is at too high a price, then there will be no winners and no capacity payments awarded. The supply bid/quantity pairs are aggregated into a supply curve, and the intersection of this with the demand curve determines the amount of new capacity added and the capacity payment that each bidder receives. We call this a multi-unit uniform price auction since the number of winners is variable but the price paid to each winner is the same, regardless of whether or not their bid was lower than the clearing price. On the bidder's side, winning the auction awards payment p and costs c, the levelized cost of that new capacity, for a return of $p - c$. Entering the auction and not winning costs nothing since the obligation to provide capacity is not present. In this paper we do not consider other costs from entering the auction such as administrative costs.

The goal for the holder of the auction is to acquire capacity to ensure the reliability of their system at the lowest possible price. It is not to acquire as much capacity as possible or to acquire it as cheaply as possible. The demand curve set by the auction holder is essentially their willingness to pay for varying amounts of capacity, and it is possible that they "overpay" for capacity when the supply price at a given quantity is lower than the demand price at that quantity. This overpayment is less apparent when allowing for bids that are not single price/quantity pairs, but rather a full supply curve offer.

In Section 2.3 we examine the optimal behavior of a single bidder playing against the "system." We observe how her behavior changes under information uncertainty and with multi-period games. In Section 2.4 we introduce multiple bidders and competition. We start with a simple single winner auction, then extend to the multiple winners case, and end with a multiple winners, decreasing price auction. In Section 2.5 we address the welfare impacts that the models in Section 2.4 have on the bidders and the system operators. We find the average probability of winning and the average total additions to the market as a means of calculating value to bidders and system operators respectively. Total cost to the system is analyzed by weighing the cost to acquire capacity versus the cost incurred from loss of load. We finish with some extensions and conclusions in the final two sections.

2.3 Single Bidder

We first assume a single bidder in the system. We assume that the capacity payment is dependent on the other actors in the system, but exogenous in this model. That is to say, a single bidder will observe a payment p or have a prior distribution on p that reflects the behavior of all other participants in the market, but whose own bid does not affect the payment.

2.3.1 Known price

We begin by addressing the known price case and assume that the single bidder knows what the market clearing price will be. In other words, the bidder cannot affect the price, but can predict it. To ensure receipt of a capacity payment, the bidder will bid below the market clearing price p. Thus the optimal strategy in this case is to submit some bid $w \in [0, p)$. However, if the cost of entry is c, the strategy changes. The optimal bid is then

$$w(c,p) = \begin{cases} c & \text{if } c > p \\ x \in [0,p] & \text{if } c \leq p \end{cases} \tag{2.3}$$

The payoffs would be 0 if the cost c is greater than p or $p - c$ if the cost c is less than the clearing price p. So, the bidder should always bid their true cost c. If it is lower than p, she will clear the auction and receive payment p. If the bid does not clear, then she will not have to pay the cost of entry c.

There is no incentive to shade bids. Suppose $w < p < c$. This means the bid would clear the auction and the bidder would receive payment p. However $p - c < 0$ and the bidder would be worse off than if her bid had not cleared. There is also no incentive to inflate the bid. If $c < p < w$, that is the costs are lower than the price p but the bid is higher than p. This bid will not clear in the auction and there will be no payoff, whereas a bid of the true cost c provides a positive return of $p - c$. There is also the trivial case of $c < w < p$, where the bidder clears in the auction with a bid higher than their cost, and doing so confers no benefit.

2.3.2 Unknown price

Now assume the price is unknown. There is still a single bidder with cost c, and all other bidders and their behaviors are represented with a probability distribution F over the interval $[a, b]$ with $a \geq 0$ that represents the range of feasible price, so $p \in [a, b]$. For simplicity, we assume that if the generator's bid w is less than p, then she receives the payment.

The value function for this generator is a function of its bid w, and is

$$V(w) = \int_a^b \mathbf{1}_{w \leq p}(p - c) f(p) \, \mathrm{d}p. \tag{2.4}$$

The indicator function determines when the bid is less than the capacity payment. If it is, the payoff is the capacity payment minus the cost of construction $p - c$. If the bid is higher than the capacity payment, the payoff is 0.

Equation (2.4) is equivalent to

$$V(w) = \int_{w}^{b} (p - c) f(p) \, \mathrm{d}p, \tag{2.5}$$

where we have replaced the indicator function by rewriting the bounds of the integral in equation (2.4). To maximize $V(w)$, we differentiate (2.5):

$$\frac{d}{dw} \int_{w}^{b} (p - c) f(p) \, \mathrm{d}p = -(w - c) f(w) \tag{2.6}$$

It is easy to see that there is a critical point where $w = c$. We check the second order conditions and determine that the critical point occurs at a maximum, and the optimal strategy is to bid the true cost, as that maximizes the expected payoff to the bidder. Again, regardless of the distribution of F, truth-telling yields the highest expected payoff.

2.3.3 Two periods

Now consider the case where the bidder must pay the costs over the lifetime of the plant, in this case, two periods. For now we assume the distribution of prices remains the same across periods. There are two decision variables, the bid in the first period w_1 and the bid in the second period w_2. If we make the simplifying assumption that the bidder can only build capacity in the first period the value function is

$$V(w_1, w_2) = \int_{w_1}^{b} (p_1 - c) f(p_1) \, \mathrm{d}p_1 + \int_{w_1}^{b} \left[\int_{w_2}^{b} p_2 f(p_2) \, \mathrm{d}p_2 - c \right] f(p_1) \, \mathrm{d}p_1 \tag{2.7}$$

There are two components to this. The first is the expected value in the first period. The bid either clears the market or does not, and the bidder receives the payoff $p_1 - c$ or nothing. The second period payoff is conditional on the first. If the bidder cleared the auction in the first period, then the bidder is required to build and to provide the capacity, and thus must pay the costs associated with entering the market. This is why the cost is moved out of the integral in the second period. Taking derivatives we see

$$\frac{d}{dw_1}V(w_1, w_2) = -(w_1 - c)f(w_1) - \left(\int_{w_2}^{b} p_2 f(p_2)\,\mathrm{d}p_2 - c\right) f(w_1) \tag{2.8}$$

and

$$\begin{aligned}\frac{d}{dw_2}V(w_1, w_2) &= 0 - \int_{w_1}^{b} w_2 f(w_2) f(p_1)\,\mathrm{d}p_1 \\ &= -w_2 f(w_2) - \int_{w_1}^{b} f(p_1)\,\mathrm{d}p_1 \\ &= -w_2 f(w_2)\,[F(b) - F(w_1)] \\ &= -w_2 f(w_2)\,[1 - F(w_1)] \end{aligned} \tag{2.9}$$

Setting equation (2.9) to 0, there is a critical point at $w_2 = 0$. Substituting $w_2 = 0$ and setting equation (2.8) to 0 gives

$$-f(w_1)\,[(w_1 - c) + \mathbb{E}[p_2] - c] = 0$$

It is easy to confirm that at $w_1 = 2c - \mathbb{E}[p_2]$ and $w_2 = 0$, we have a maximum.

For n periods the optimal bid in the first period would be $w_1 = nc-(n-1)\bar{p}$ where $\bar{p}$ is the expected value of the capacity payment, and we define the total cost as $C = nc$. Similarly, for different distributions of prices in each of the periods we have: $w_1 = nc - \sum_{i=2}^{n} \mathbb{E}\left[p_i\right]$ and the optimal bid in the first period is still the total cost minus the sum the of expected payments from future auctions. However, for very large total cost C or low $\bar{p}$, a large bid w_1 could result in a project that is never constructed since it may never clear in the auction.

Thus the optimal bidding strategy for n periods is

$$w_t = \begin{cases} nc - \sum_{i=2}^{n} \mathbb{E}[p_i] & \text{if } t = 1 \\ 0 & \text{otherwise} \end{cases}$$

Historical bids from PJM auction data support this strategy. The majority of existing capacity entered in the auction is bid in at a price of \$0 because once a plant is built, the generator wants its bid to clear in every subsequent capacity auction.

2.4 Multiple Bidders

2.4.1 Single winner

We now introduce competition into the model. Each bidder i has some cost c_i of entry, which follows a common distribution $C_i \sim F(\cdot)$. The costs are independent. Again there is a capacity payment p and a single winner who will receive it if they are (1) the lowest bidder, and (2) bid lower than payment p. Each bidders' bid is some function of their cost. Assume that each bidder $j \neq i$ uses the same bidding strategy $w_j = w(c_j)$. We will ignore asymmetric equilibrium in this model.

The value function of bidder i is then

$$V_i(c_i, w_i) = (p - c_i) \cdot \Pr\left[w_j > w_i,\, \forall j \neq i\right] \cdot \mathbf{1}_{w_i \leq p} \tag{2.10}$$

This value function represents the expected value of the bid w_i and the cost c_i. The payoff the bidder receives is equal to the capacity payment received less the cost if they win, or zero otherwise. Since there is no obligation to provide capacity if the bidder does not win, there will be no costs incurred. Again, we ignore administrative costs.

Bidder i wants to choose w_i such that

$$\max_{w_i} (p - c_i) \left(1 - F\left(w^{-1}(w_i)\right)\right)^{n-1} \cdot \mathbf{1}_{w_i \leq p} \tag{2.11}$$

Calculating the distribution of bids would require transforming the distributions, which is why in Equation (2.11) we transform the probability statement $\Pr[w_j > w_i]$ into $\Pr[c_j > w^{-1}(w_i)]$, for which the distribution is assumed known.

In this example, since the bidders know p, and they will know their c_i, then w_i needs to be either 0 or greater than p. If the payment is higher than the cost ($p > c_i$), then one should bid as low as possible to beat out the other bidders and win the auction. (Here there is bid shading since the actual cost of entry will almost certainly be non-zero which is different from the case where there was a single bidder acting alone.) If the payment is lower than the cost, then a bidder should bid higher than p to avoid winning the auction.

Now let us assume that all bidders have a common prior distribution $P \sim H(\cdot)$ on the payment. The expected payoff becomes:

$$V_i(c_i, w_i) = \int (p - c_i) \cdot \Pr[w_j > w_i,\ \forall j \neq i] \cdot h(p) \cdot \mathbf{1}_{w_i < p} \, \mathrm{d}p \tag{2.12}$$

Bidder i wants to choose w_i such that it maximizes

$$\max_{w_i} V_i(c_i, w_i) = \int_{w_i}^{b} (p - c_i) \left(1 - F\left(w^{-1}(w_i)\right)\right)^{n-1} h(p) \, \mathrm{d}p \tag{2.13}$$

The first order condition is

$$\begin{aligned} &-(w_i - c_i) h(w_i) \left(1 - F\left(w^{-1}(w_i)\right)\right)^{n-1} \\ &\quad + \int_{w_i}^{b} (p - c_i)(n-1) \left(1 - F\left(w^{-1}(w_i)\right)\right)^{n-2} \frac{-f\left(w^{-1}(w_i)\right)}{w'\left(w^{-1}(w_i)\right)} h(p) \, \mathrm{d}p = 0 \end{aligned}$$

Replacing w_i with $w(c)$ and dropping subscripts we have

$$-w'(c)(w(c)-c)h(w(c)) = (n-1)\frac{f(c)}{1-F(c)}\int_{w(c)}^{b}(p-c)h(p)\,\mathrm{d}p. \tag{2.14}$$

Equation (2.14) will be difficult to solve without making some additional assumptions. For simplicity, we assume each bidder has some cost c_i drawn independently from a uniform distribution $U(0,2)$ and the prior distribution on the price p is also a uniform distribution $U(0,2)$. In the context of a capacity auction for electricity, a cost of 1 would be equivalent to one times the cost of new entry (CONE), a commonly used value that represents the levelized cost of a new combustion turbine plant, usually reported in dollars per megawatt-year (around $121,000/MW-year (PJM (2014)). Thus, in this example, the possible costs of bidders and the possible capacity payments both range from zero to twice the cost of new entry, and their means are equal to the cost of new entry.

The optimal bid differential equation then becomes

$$w'(c) = -\frac{(n-1)(w(c)-2)(2c-w(c)-2)}{2(c-2)(c-w(c))}$$

The solution to this differential equation with boundary condition $w(2)=2$ is:

$$w(c) = \begin{cases} \frac{2(nc-(n-1))}{n+1} & \text{if } c > (n-1)/n \\ 0 & \text{if } c \leq (n-1)/n \end{cases}$$

In actuality, the solution is not a piecewise function; we have truncated the lower end of the range at 0 since negative bids are not allowed. Figure 2.2 shows the optimal bids as functions of cost for various numbers of bidders n and a gray dashed line representing truthful bidding where $w(c)=c$. We observe bid shading at every realization of cost; i.e. every bid function lies below the truthful bidding line. Regardless of whether your cost is above or below the cost of new entry, you will bid less than your cost (with the exception of a single point at the upper limit of the cost, $w(2) = 2$). Additionally, the amount by which you shade your bid decreases as your cost increases. For $n = 2$ bidders, a cost of $c = 0.75$ (or \$90,750/MW-year), the optimal bid is $w = 0.33$ (or \$40,333/MW-year), a 55% decrease. A cost of $c = 1.5$ results in a bid of $w = 1.33$, only an 11% decrease.

When there is more competition, or bidders, the severity of the bid shading increases. For example, with two bidders, and a cost equal to the cost of new entry $c = 1$ or (\$121,000/MW-year), the optimal strategy would be to bid $w = 0.66$. With $n = 3$ bidders, the same cost results in a bid of 0.5, and with $n = 4$, a bid of $w = 0.4$. There are two probabilities acting as opposing forces, the probability that you beat out the other bidders making you want to lower your bid, and the probability that you clear in the auction but at a clearing price that is lower than your true cost c, suffering a loss.

Another way to look at this is to say that the region of 0 bidding increases as n increases, since the probability that a competitor's bid c_j will be lower than the bidder's cost c_i increases with n. Therefore one bids 0 to ensure that one has the lowest offered price. It should be noted that this region of 0 bidding will always be some region $c \in [0, t]$ where $t \leq 1$. This is because the expected value of the payment is equal to 1. By bidding 0 when your true cost is below 1, your expected gain is still positive.

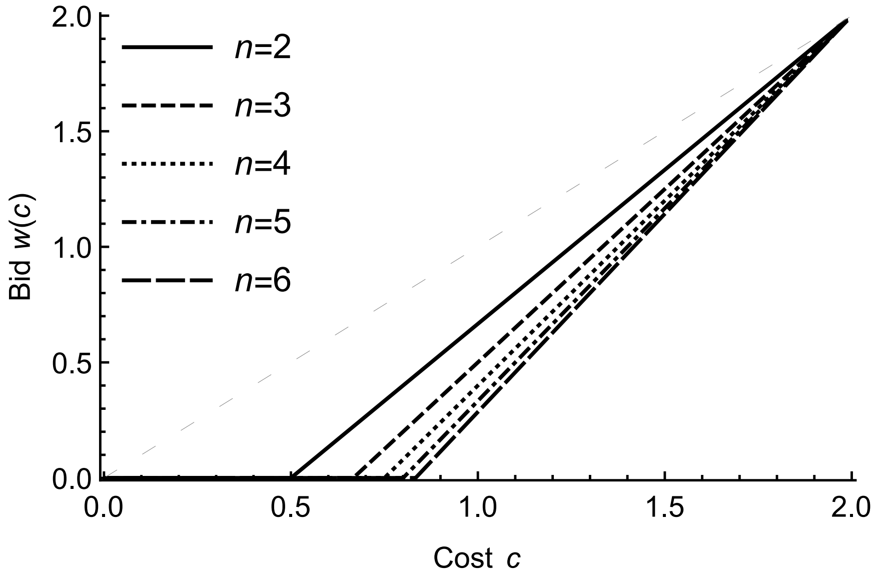


Figure 2.2: Optimal bid functions in the single winner case. Note the upper most gray line on the diagonal represents truthful bidding, e.g. $w(c) = c$.

In Appendix 1 we analyze a case with a less "convenient" choice of bounds on the uniform distribution to show that the result still generalizes.

2.4.2 Multiple Winners

2.4.2.1 Fixed price

We extend the analysis to have $k > 1$ winners, but still a fixed price unaffected by the number of bidders. We can consider this as a proxy for elastic or nearly elastic demand curves or minuscule supply additions such that the total added capacity does not change the capacity payment to winners. PJM's earlier capacity credit market was similar to this (Sener and Kimball (2007)). Capacity up to a certain level was paid a fixed price, and any capacity over that benchmark received nothing.

In this model, rather than needing to be the lowest bidder, a winning bidder only needs to (1) bid lower than the payoff, and (2) be one of the k lowest bidders. The value function for a bidder is

$$V_i(c_i, w_i) = \int (p - c_i) \cdot \Pr\left[W_{(k)} > w_i\}\right] h(p) \cdot \mathbf{1}_{w_i<p}\, \mathrm{d}p. \tag{2.15}$$

We assume that all bidders share a prior distribution $h(p)$ on the payoff. We define $W_{(k)}$ to be the kth order statistic of the $n-1$ other bids. Because we are looking for a symmetric equilibrium, and the bidding function should be monotone, we can rewrite this in terms of the order statistic of unknown costs C, for which a distribution can be easily found:

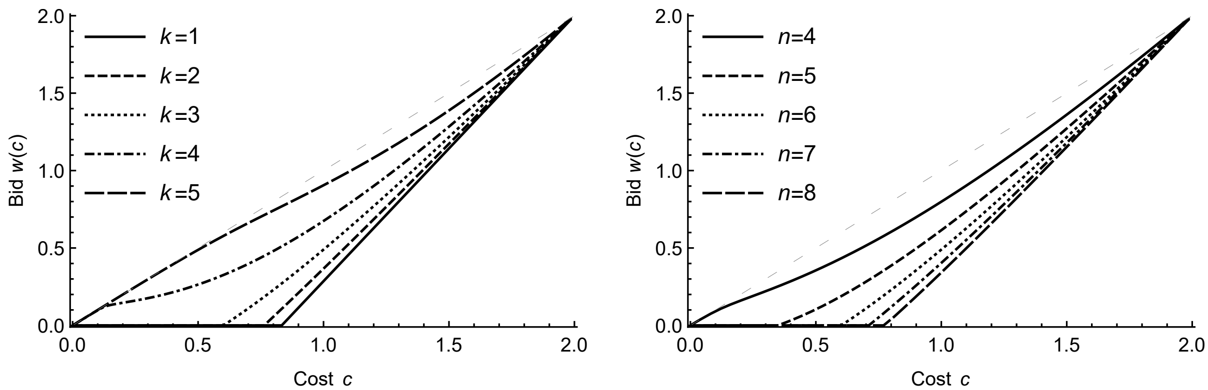
$$V_i(c_i, w_i) = \int (p - c_i) \cdot \Pr\left[C_{(k)} > w^{-1}(w_i)\}\right] h(p) \cdot \mathbf{1}_{w_i<p}\, \mathrm{d}p \tag{2.16}$$

If we assume again that the distribution of costs and the common prior on p follow a $U(0, 2)$ distribution, we obtain the following first order condition

$$w'(c) = \frac{(w-2)c^{k-1}(2c - w - 2)(2-c)^{n-1-k}}{2^n(c-w)B(k, n-k)(1 - I_{\frac{c}{2}}(k, n-k))} \tag{2.17}$$

where B is the beta function and $I_{\frac{c}{2}}$ is the regularized incomplete beta function $(B(x; a, b)/(B(a, b))$. A closed form solution cannot be found, so the differential equation solver LSODA was used. The shapes of the optimal bid functions can be seen in Figure 2.3. Note that in Figure 2.3a, for a fixed number of bidders $n = 5$, as the number of winners k increases the optimal bid function approaches $w(c) = c$. Because the likelihood of clearing the auction increases with the number of accepted winners, one does not have to submit such lowball bids in an effort to beat out other bidders. However, the function alternates between convex and non-convex due to large portion of almost surely 0 probability in the Beta distribution when n and k are close.

For a fixed number of winners k, we see that bidders have the incentive to bid a smaller and smaller fraction of their true costs as competition increases. As n increases the 0 region of bids will increase up to 1 times CONE, as costs below 1 have a positive expected profit. This is very reminiscent of the single winner case. Figure 2.3b shows that misrepresentation of costs is less of an issue as n approaches $k = 3$; the intensity of competition decreases as n approaches k. With more chances to win a payment, having a bid that more accurately represents the true cost minimizes the chance of a net loss.



(a) Optimal bid function for various k, $n = 6$. (b) Optimal bid function for various n, $k = 3$.

Figure 2.3: Optimal bid functions in multiple winner, uniform price auction. Note the upper most gray line on the diagonal represents truthful bidding, e.g. $w(c) = c$.

2.4.2.2 Decreasing demand curve

The previous section dealt with the analysis of a typical auction with a constant payout. Now we extend the game to have a uniform payout as a function of the offers. The more winners the lower the payout. We approximate the demand curve with a linear function with slope α. Since all suppliers are identical, each unit supplied will move along the demand curve an equal amount. We simplify this by fixing some base price $p_1 = p$ for 1 accepted unit, and $p_{i+1} = p - i\alpha$ for additional cleared supply.

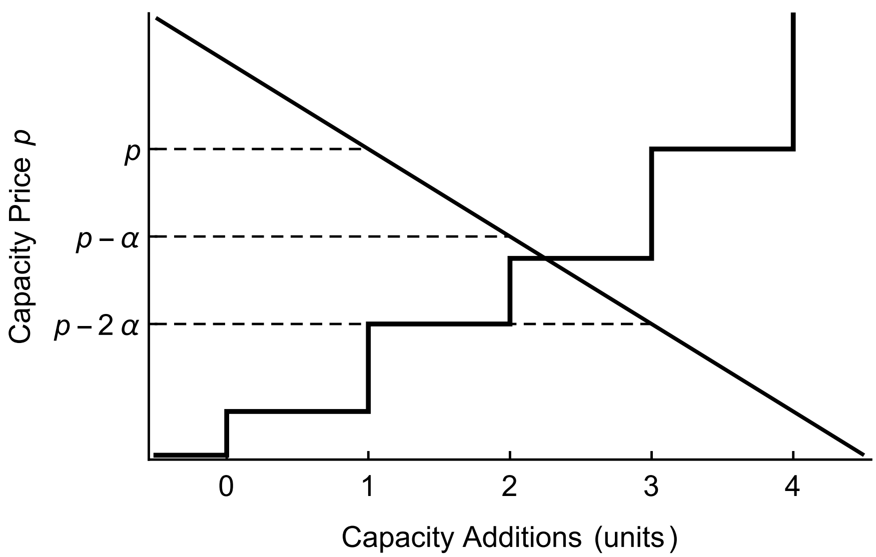


Figure 2.4: Supply and demand curves

Some examples will help illustrate the auction clearing mechanism. In Figure 2.4, we have a demand curve and a supply curve created from sorted and stacked bids. The price for a single unit of capacity is p; for two, $p - \alpha$; for three, $p - 2\alpha$ and so on. The supply curve intersects the demand curve between two and three units; because we do not allow fractional capacity additions only two units clear, and they clear at a price of $p - \alpha$, even though the supply costs less. Table 2.1 has various possible scenarios of bids and outcomes for three bidders X_1, X_2, and X_3, and their respective bids W_1, W_2, and W_3.

Let $p_1 = 1.6$, $\alpha = 0.4$, $p_2 = 1.2$, $p_3 = 0.8$, etc.

Note that in the second row of Table 2.1, though X_2 and X_3 are less than p_1, only the lower bid W_3 receives the payoff. For n suppliers to receive a payment, those n bids must

W_1	W_2	W_3	Outcome
1.10	0.60	1.60	Two winners, X_1 and X_2, receiving $p_2 = 1.2$.
1.80	1.50	1.40	One winner, X_3, receiving $p_1 = 1.6$.
0.40	0.80	0.60	Three winners, X_1, X_2, and X_3 , receiving $p_3 = 0.8$.
1.74	1.80	1.62	No winners.

Table 2.1: Example payments to bidders.

all be less than or equal to p_n.

If we repeat our assumption that bidders are identical, and that the price is random according to some distribution $P \sim H(\cdot)$, our value function is now much more complicated:

$$
\begin{aligned}
V_i(c_i, w_i) = &\int (p - c_i) \cdot \Pr\left[W_{(1)} > p - \alpha\}\right] h(p) \cdot \mathbf{1}_{w_i < p-\alpha} \, \mathrm{d}p \\
&+ \int (p - c_i) \cdot \Pr\left[W_{(1)} > w_i\}\right] h(p) \cdot \mathbf{1}_{p-\alpha < w_i < p} \, \mathrm{d}p \\
&+ \int (p - \alpha - c_i) \cdot \Pr\left[W_{(1)} \leq p - \alpha \cap W_{(2)} > p - 2\alpha\right] h(p) \cdot \mathbf{1}_{w_i < p-2\alpha} \, \mathrm{d}p \\
&+ \int (p - \alpha - c_i) \cdot \Pr\left[W_{(1)} \leq p - \alpha \cap W_{(2)} > w_i\right] h(p) \cdot \mathbf{1}_{p-2\alpha < w_i < p-\alpha} \, \mathrm{d}p \\
&+ \int (p - 2\alpha - c_i) \cdot \Pr\left[W_{(2)} \leq p - 2\alpha \cap W_{(3)} > p - 3\alpha\right] h(p) \cdot \mathbf{1}_{w_i < p-3\alpha} \, \mathrm{d}p \\
&+ \int (p - 2\alpha - c_i) \cdot \Pr\left[W_{(2)} \leq p - 2\alpha \cap W_{(3)} > w_i\right] h(p) \cdot \mathbf{1}_{p-3\alpha < w_i < p-2\alpha} \, \mathrm{d}p \\
&+ \\
&\vdots \\
&+ \int (p - n\alpha - c_i) \cdot \Pr\left[W_{(n)} \leq p - n\alpha\right] h(p) \cdot \mathbf{1}_{w_i < p-n\alpha} \, \mathrm{d}p
\end{aligned}
$$

where $W_{(j)}$ is the order statistic of the $n-1$ other bids. Whereas before we only compared bids against other bids, here we must also compare bids against unknown prices, which depend on offers of other market participants. As seen in the example, there are two ways for a bidder to realize a price p_t: either ensure that the $t-1$ lowest competitors' bids are less than p_t and the others are bids higher than p_{t+1} if her bid is less than p_{t+1}, or that the $t-1$ lowest competitors' bids are less than p_t and the other bids are higher than her bid if her bid is less than p_t.

For n bidders and a demand curve slope $-\alpha$, we obtain the differential equation

$$w'(c) = \frac{n-1}{4}\left(\frac{\alpha^2 + 2\alpha(w(c) - c)}{c - w(c)}\right), \tag{2.18}$$

which has solution

$$w(c) = \frac{a^2(-(n-2)) + 2ac(n-1) + 4c}{2a(n-1)+4}$$

when $w(2) = 2 - \frac{a^2(n-2)}{2a(n-1)+4}$. Unlike the multiple winners uniform price auction, the optimal bid function again has the kinked shape from the single winner auction seen in Figure 2.2. Figure 2.5 shows optimal bid functions for various bidders with $\alpha = 0.3$. It is not as apparent in the figure, but for lager α levels there is more separation between the bid functions for various n and higher n values shift the x-intercept to the right. The $n = 2$ line is the topmost.

This α value represents a demand curve steeper than that used by PJM and there is not much sensitivity to a change in the number of bidders. As the number of bidders increases, the bid function shifts rightward along the x-axis. As α increases, there is a larger impact on the shift of the curve and there variability in the prices increases. The severity of the bid shading is reduced, since each bidder is now much more likely to receive a low payment, even if the starting price p is high.

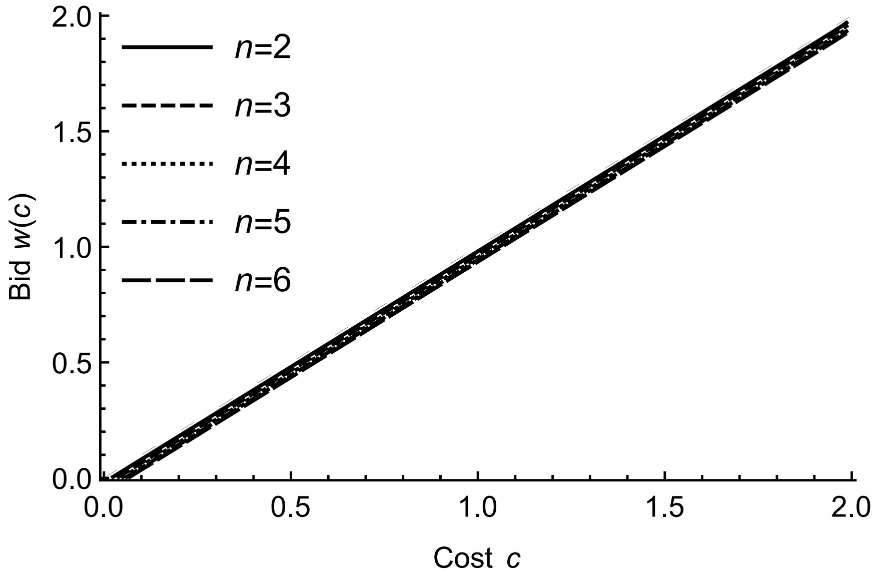


Figure 2.5: Descending price auction with $\alpha = 0.3$.

2.5 Welfare

We now study the impacts of this auction on the welfare of the bidders and the auction holders, the ISOs. What separates this auction from many other auctions is the inclusion of a demand curve that is functioning similarly to a reserve price. Every capacity payment is willingly paid by the ISO, as set by their demand curve, to add additional megawatts of capacity. The results produced in this section are obtained from simulation where exact results could not be computed, especially in the multiple winner cases. Costs are drawn from a Uniform$(0, 2)$ distribution, and the optimal bid function is applied to these costs for every bidder. Ties are broken randomly. Clearing prices and payoffs are calculated and 200,000 iterations were run to obtain convergence.

We begin this section by discussing the simplest cast, where there are n entrants and one winner, and the price paid is randomly drawn from a Uniform(0,2) distribution. We then extend this case to have k winners. The price paid is still randomly drawn from the Uniform(0,2) distribution, but now up to k units that clear get that price. Lastly we examine the case where an initial price for one unit of generation is drawn from the Uniform(0,2) distribution, but the price paid decreases as a function of additional generation. We conclude the section with a comparison example of the multiple winners, flat price case and multiple winners, decreasing demand curve case. The total costs to the system are broken down in terms of costs due to potential outages, and costs of acquiring new capacity.

In all scenarios, the auction is inefficient. An efficient auction is one where the winners are those with the highest valuations, or in this case, the ones with the lowest costs. Because there are regions of zero bids in each optimal bid function, there can be a situation where multiple bidders bid zero. In this case, the winner would be determined randomly, which could result in a bidder with a higher true cost winning the auction.

2.5.1 One winner

For the one winner case, the bidders have an incentive to misrepresent their costs. We can calculate the average winning bid in the single winner case.

$$\int w \, \mathrm{d}W_1(c) = \int_{\frac{n-1}{n}}^{1} \frac{nc-(n-1)}{n+1} n \left(1-\frac{c}{2}\right)^{n-1} \mathrm{d}c = \frac{\left(\frac{1}{n}+1\right)^n}{2^{n-1}(n+1)} \tag{2.19}$$

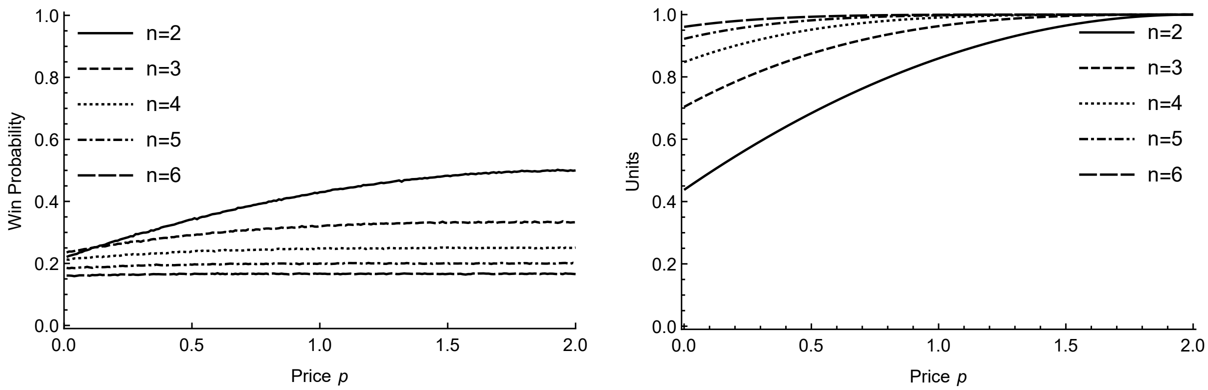
where n is an integer and greater than 1. As the number of participants increases, the average winning bid approaches zero. Equation 2.19 has two intuitive implications as the number of participants increases: (i) the higher the probability that some participant has a low cost; (ii) the optimal bid function has a larger region of 0 bids. Because the winning bid goes to zero, for any given n the probability of having a unit of capacity accepted approaches 1 since the actual payment p is in the interval $[0, 2]$.

The probability that the ISO receives new capacity at a given price is equal to

$$\Pr\left[C_{(1)} < w^{-1}(p)\right] = F_n\left(\frac{(n+1)p + 2(n-1)}{4n}\right). \tag{2.20}$$

In Equation 2.20, F_n is the CDF of the beta distribution with parameters $(1, n)$. As the price increases, the probability that the ISO acquires a unit increases. In the two player case, the average probability that a bid gets accepted is 50% when the price is 0, and goes to 1 as the price goes to 1. As n increases, for any given price p, and $\{n_1, n_2 \,|\, n_2 > n_1\}$, there is a higher chance of accepting a unit at any given price level, that is $F_{n_2}(w_{n_2}^{-1}(p)) > F_{n_1}(w_{n_1}^{-1}(p))$. This is a result of the inverse bid functions w^{-1} increasing in n and $\mathrm{Beta}(1, s)$ stochastically dominating $\mathrm{Beta}(1, t)$ for $1 \leq s < t$. Figure 2.6 shows the average probability of winning for an individual and the mean number of units accepted in the auction at varying realizations of p.

As more participants enter the market, each individual bidder becomes less likely to win (see Figure 2.6a), but the ISO is more likely to come away with a unit of generation (see Figure 2.6b). It is in the auction holder's best interests to encourage more participation as this results in greater overall system stability from the acquisition of new capacity. Conversely, the generator would prefer fewer participants since this increases the chance of winning. However, because entry is costless and the cost c is only incurred if the bidder wins, there is little downside to entering the auction in these situations. There is a balance between the ISO, who wants to obtain capacity, and thus a higher number of entrants, whereas the generators want to clear in the auction and thus a lower number of entrants.



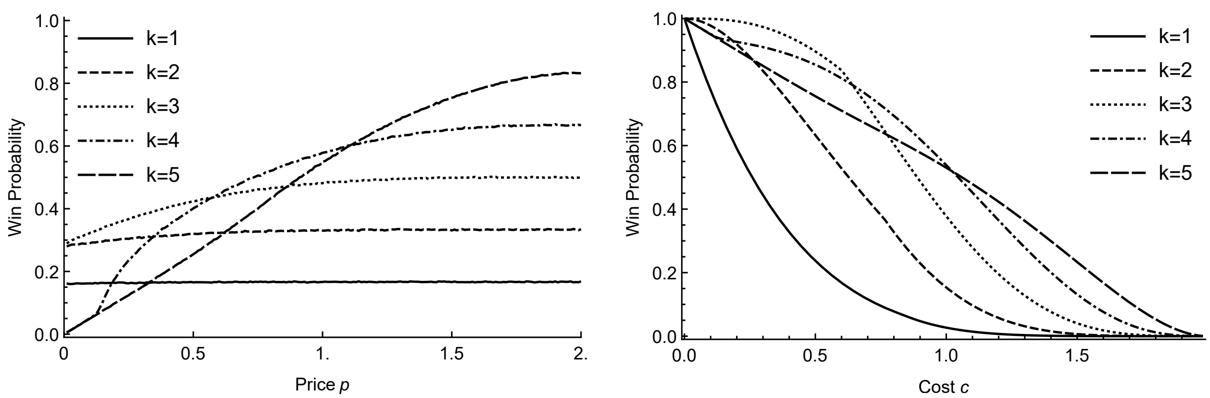
(a) Individual probability of winning (b) Mean number of units accepted

Figure 2.6: Single winner.

2.5.2 Multiple winners, flat price

As in the single winner case, when there are multiple winners the average winning bids decrease as the number of participants n increases. That is to say, the average lowest bid in the two bidder case is higher than the average lowest bid in the ten bidder case. This is true for all the accepted bids. However, the highest bid that gets accepted is increased because there are more bidders.

The number of bidders clearing in the auction increases as the price increases, as expected. The higher the price, the more likely it is that there are k winners below that price. If the clearing price is low there might be less than k bidders accepted, even though there are up to k winners. Figure 2.7 is a graph of a bidder's probability of winning. The interesting point here is that the probabilities are not monotonic in k. Recall in Figure 2.3a that the bid function begins with the kinked shape similar to the single winner case, but approaches the more linear shape as k approaches n. Thus at low values of p, even though there can be more winners, the bid function has changed such that the chance of submitting a bid below p is reduced. For high values of p we see the expected probability of winning is k/n in Figure 2.7a. This is much like the win probability in the single winner case from the previous section, except the asymptotes were at $1/n$. Another interesting byproduct of the unconventional bid functions is the win probability as a function of cost seen in Figure 2.7b. At costs greater than 1 the probability of winning increases in k, but no such generalizations can be made at the lower costs.



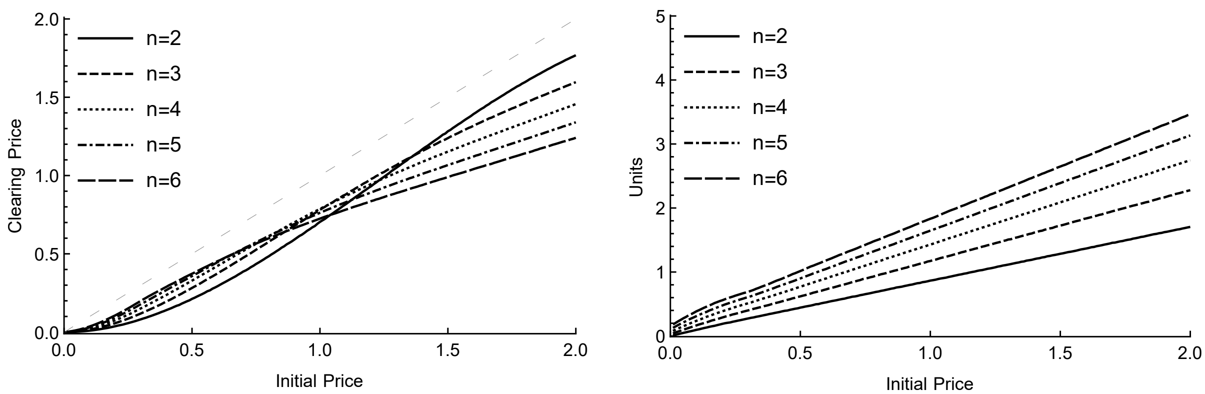
(a) Win probability against price

(b) Win probability against cost

Figure 2.7: Individual probability of winning, uniform price with five bidders, ($n = 6$).

2.5.3 Multiple winners, decreasing demand curve

For the descending demand curve case, Figure 2.8a shows the average clearing price in the auction when p_1 starts at the level given on the x-axis. Note that the clearing price can never be higher than the initial price, indicated by the dashed line $y = x$. With more bidders, there are more chances for cheaper generators to enter, thus driving down the price. At $p_1 = 2$, $\alpha = 0.3$, and $n = 6$, the average clearing price is a little under 1.2 as seen in Figure 2.8a. This indicates that on average, a little over 3 units were accepted at that price as corroborated by the output in Figure 2.8b. The average number of additions of units of capacity increases as n increases, but it should be noted that the proportion of winners to entrants decreases. These values are dependent on the α level; a steeper demand curve would mean fewer additions of units of capacity and vice-versa. The amount of new additions is heavily limited by α. At $p_1 = 2$ and $\alpha = 0.4$, the possible prices are $\{2.0, 1.6, 1.2, 0.8, 0.4, 0\}$. This means that for more than 5 bidders to be accepted, all entrants must be bidding in a cost of 0.

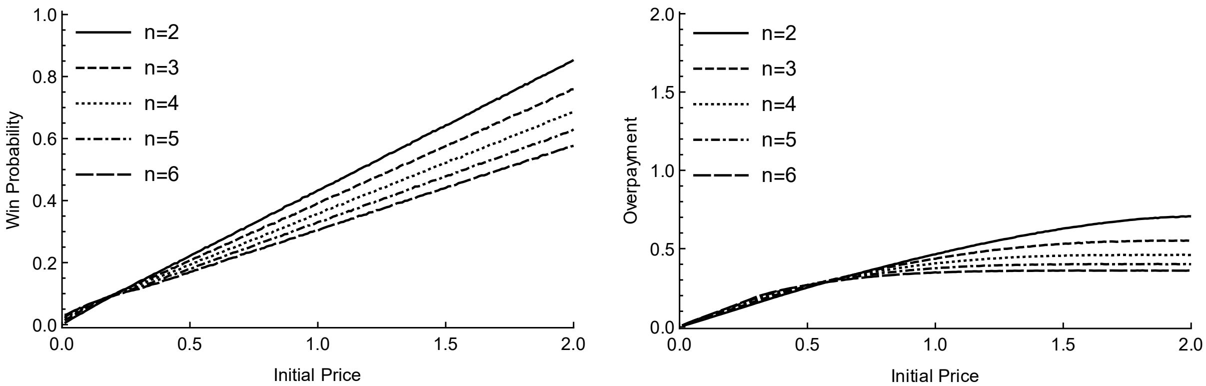


(a) Clearing price

(b) Mean number of units accepted

Figure 2.8: Demand curve based auction, $\alpha = 0.3$.

As expected, when the number of participants increases, the individual probability of winning decreases. We see this in Figure 2.9a. When the auction is settled, the price from the demand curve may be above the highest clearing bid. This results in an overpayment to the marginal producer (an example can be seen in Figure 2.4). This is less of an issue in a situation where bidders can submit piecewise cost curves rather than flat prices or when fractional capacity can be accepted. In Figure 2.9b we see that there is a limit to the total over-payment. We define over-payment as (clearing price - marginal bidder) $\times$ total units accepted. From earlier, we know that more bidders increases the amount of accepted capacity (Figure 2.8b) and the total cost of paying the winners. However, the increase in bidders also creates more competition and tightens the gap between the demand curve and the marginal bidder.



(a) Individual probability of winning auction

(b) Total overpayment

Figure 2.9: Demand curve based auction, $\alpha = 0.3$.

2.5.4 Comparison Example

Let us construct a simple example to examine the effectiveness of different types of auctions. We can determine the total cost to the system by adding the capacity payments for procuring capacity and the expected cost due to loss of load events (LOLE). Using data from PJM, we fit a curve to model the relationship between the installed reserve margin (IRM) and the LOLE. The installed reserve margin is set by PJM so that the loss of load probability (LOLP) is 1 day in 10 years. At a level of 3% below the target reserve margin, the LOLP is about 4 days in 10 years, and at a level of 5% above, the LOLP is a little under 1 day in 100 years. 1 day in 10 years is equivalent to 0.1 days in a single year, 5 days in 10 years is 0.5 days in a year, and so on. The relationship is closely approximated by the equation LOLP $= 0.1011 \times \text{IRM}^{-49.01}$ as shown in Figure 2.10.

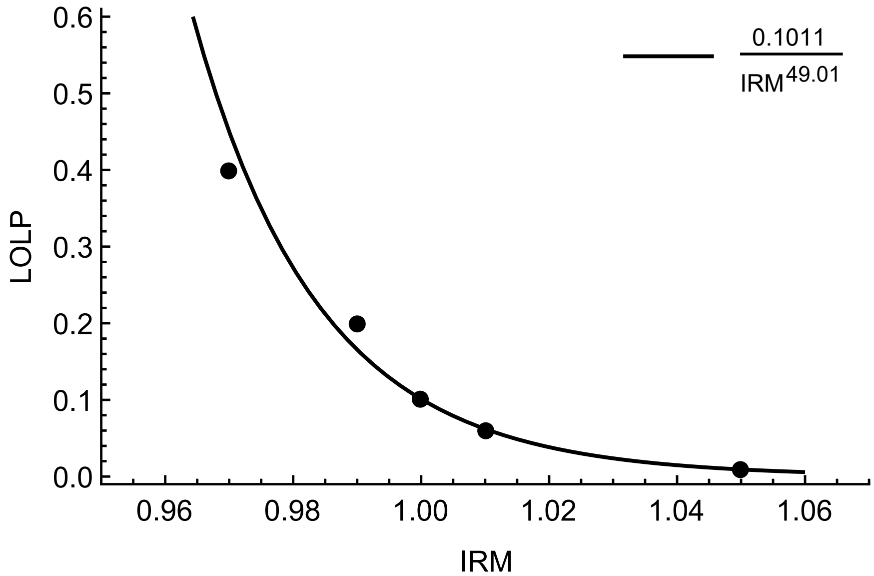
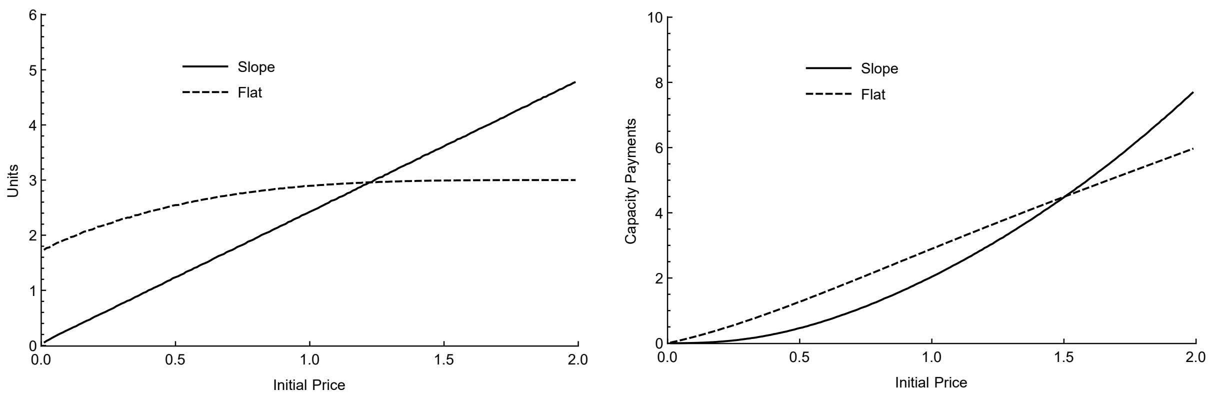


Figure 2.10: Regression to determine relation between IRM and LOLP

Assume that the market reserve margin currently sits at 97% of the IRM, and that each new unit would add 1% of the desired reserve margin, so to reach the desired LOLP of 1 day in 10 years (1-in-10), three units would need to be accepted. In the flat auction, we would choose three winners from n bidders, and in the decreasing demand curve case, the slope is set to to $\alpha = 0.1$ to most closely resemble PJM's demand curve. Here, we set $n = 6$ bidders. The results can be seen in Figures 2.11 and 2.12. The lowest lines represent a value of lost load (VOLL) of \$5,000/MWh, a common value used by ISOs (Frayer et al. (2013)). We calculate the total cost to the system as VOLL $\times$ LOLP $+$ units clearing $\times\, p$ where p is the capacity payment. For a 1-in-10 LOLP, this is equivalent to a VOLL cost of \$12,000/MW-year (=\$5,000/MWh $\times$ 24 h/day $\times$ 0.1 day/year). Compared to a net cost of new entry of approximately \$121,000/MW-year (PJM (2014)) it is easy to see that it is cheaper to have a higher probability of loss of load and less capacity than it is to pay for new capacity to reduce the loss of load probability. Each additional percentage increase in installed reserve margin has diminishing returns on the reduction in LOLP based on the model fit earlier.



(a) Average units accepted (b) Expected cost from capacity payments

Figure 2.11: Comparison of auction formats with IRM = 97%, $\alpha = 0.1$, and $n = 6$.

The number of units clearing and the capacity payments to generators are unaffected by the VOLL, which is why there are only two curves in Figure 2.11. Looking at the results of the simulation, we turn our attention to Figure 2.11a. This figure shows the average units accepted in both the flat price and sloped demand curve cases assuming an initial clearing price. In the flat case, the initial clearing price is the price that is given to all units that clear. Because of the hockey stick shape of the bidding functions in the flat-price case, a significant amount of capacity clears at low prices, almost two units when the capacity payment is zero. Conversely, in the sloped demand curve case, an initial price of zero means that for any number of units to clear they must be bid in at a price of zero, which is much less likely given the optimal bid function for sloped demand curves. Because we started at

97% of the installed reserve margin and we seek to reach 100%, in the flat price case only three units(=(100 − 97)%/1%) will be accepted, which is why the units clearing tapers off at three in Figure 2.11a. All units could clear in the sloped case which is why the average number of units accepted increases up to five at the highest initial price of 2.

Figure 2.11b shows the expected cost of capacity payments in the two cases. The capacity payment is equal to the number of units clearing times the final clearing price. Again, in the flat price case the final price is the same as the initial price, but in the sloped case it depends on the number of units clearing. If five units clear in the sloped case at an auction that started with the initial price at 2, then the final price is $2 - 4 \times \alpha = 1.6$. The capacity payment would then be 8 total for the ISO. In contrast with the flat case, at a price of 2, only 3 units are being accepted, resulting in a lower total capacity payment of 6. For low initial prices, the sloped case results in lower total capacity payments in comparison to the flat price case due to the fewer number of units accepted. As initial price increases, the total payment curves cross and the sloped case becomes more expensive. Even though the capacity price is lower in the sloped price case than the flat price case, the more units accepted increases the total payment by the ISO to generators.

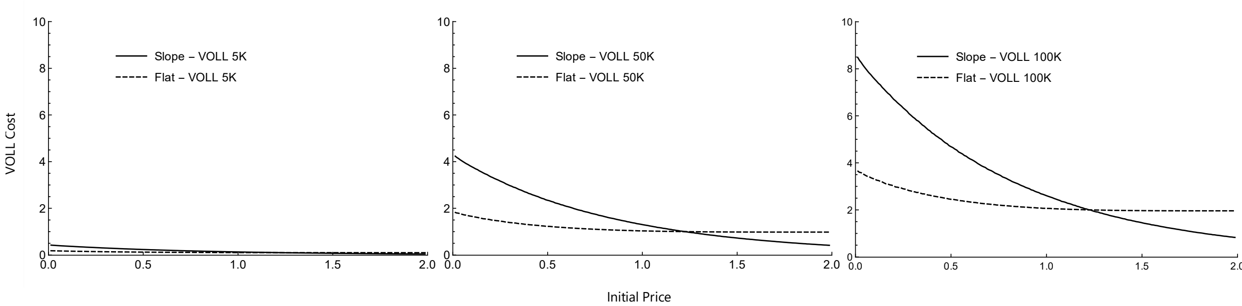


Figure 2.12: Expected VOLL costs with IRM = 97%, $\alpha = 0.1$, and $n = 6$

Looking now at the costs as a result of lost load, we see in Figure 2.12 that expected VOLL costs decrease as initial price increases in all cases and scenarios. Recall from Figure 2.11a that the number of units clearing was increasing as a function of initial price. The formula for VOLL costs is VOLL $\times$ LOLP, and since VOLL is fixed the only thing affecting this cost is the LOLP, which is a function of how many units clear in the auction. The more units that clear, the higher the installed reserve margin, and the lower the loss of load probability. Again, at low prices, because few units clear in the sloped price case, the VOLL cost is very high when compared to the flat price case for any given level of VOLL. However, as initial price rises, more units clear in the sloped case, while the flat price case is capped at three units, allowing the IRM to increase beyond 100%. This in turn results in a lower LOLP which is why the curves cross and even leads to the sloped price case with a 100K VOLL having a lower VOLL cost than the flat case with only a 50K VOLL.

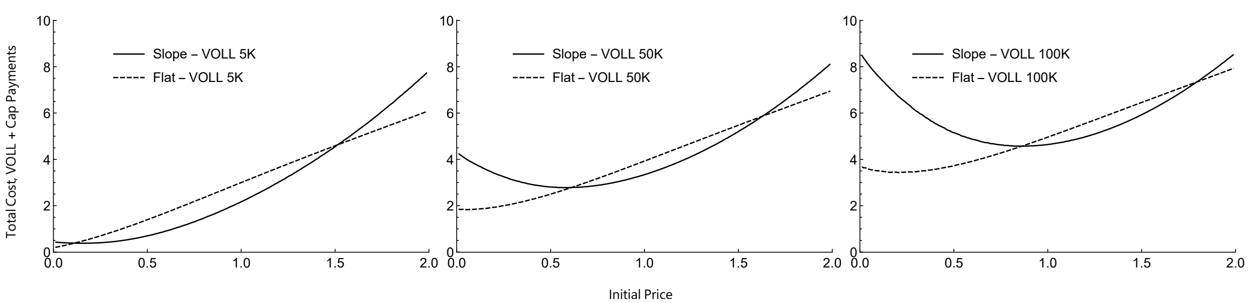


Figure 2.13: Total expected costs with IRM $= 97\%$, $\alpha = 0.1$, and $n = 6$

Combining the VOLL cost and capacity cost graphs, we have Figure 2.13. At low prices, the capacity payments to generators were low in the sloped case, but this is offset by the high VOLL costs at these low prices. The rapid decrease in the VOLL costs in the sloped case cannot overcome the increasing capacity payments as initial price increases resulting in the subsequent dip and then rise in the sloped price case. Contrasted with the flat case which has a brief dip at low prices and then a steady rise as we reach the threshold for having three units clear in the auction. It is interesting to note that the range of prices at which the sloped case results in lower total costs for a given level of VOLL decreases as VOLL increases. At large values of VOLL it would never be optimal to use a sloped curve auction. However, the initial prices cannot be guaranteed, so an average was taken for comparison's sake (Table 2.2). For a 5K VOLL, the sloped case is cheaper on average, but ends up more expensive at greater values of VOLL. The high costs at the fringes of the domain of initial prices outweighs the savings in the middle. At higher initial IRMs, the weaknesses of the sloped case at low prices would disappear, as the penalty from VOLL cost for having a low IRM is greatly reduced. Further analysis would be needed to examine all possibilities and scenarios.

VOLL ($/MWh)	Sloped	Flat
5,000	2.79	3.01
50,000	4.26	4.04
100,000	5.88	5.18

Table 2.2: Average total expected costs over possible initial prices.

For the commonly accepted value of lost load of 5,000 $/MWh, the sloped case is on average a cheaper means of procuring reliability, or a little under 8% less than the flat case. Since in the end the costs are passed on to the consumers, such an auction scheme would be preferable as it would result in lower electricity prices. However, the true VOLL can be much higher for certain customers Frayer et al. (2013) and in this situation the flat case average cost is 12% lower than the sloped case when the VOLL is 100,000 $/MWh. Many ISOs have regulations in place that don't allow generators to report anything but their true costs, thus the optimal bidding behavior would not be possible. The sloped demand curve results in nearly truthful bidding, and would perhaps be beneficial to ISOs by reducing resources dedicated to compliance.

2.6 Extensions

The models presented in this paper provide a basic understanding of the behaviors of participants in a capacity market. However, outside the scope of this paper are many extensions to the basic model which would add value. The most relevant may be that of asymmetric bidders. In an actual capacity market, many different types of capacity are being bid in; e.g. coal, natural gas, nuclear, etc. These different types have different costs and also different capacities. Relaxing the identical bidders assumption would more clearly reflect actual market conditions. Similarly, the choice of a prior distribution or the incorporation of Bayesian updating in a repeated auction framework would also shed light on generator behavior. The capacity price that settles in year t could be used to obtain a prior distribution to use in year $t+1$, affecting a generator's decision to bid in a given year or to delay to the next.

Bidders were treated as risk neutral, and thus were content to shave their bids since the probability of the price being greater than their costs was $1-c$. The assumption of risk averse utility functions for the bidders would better reflect observed market participant behavior, and would certainly impact the degree that bidders under report their cost.

Most demand curves used in these auctions are not linear. They are often piecewise functions with differing slopes at differing levels of installed capacity, reflecting the need for additional capacity. Recently PJM has discussed changing the shape of their demand curve (it currently is concave, a feature not often seen in demand curves) to a more traditional convex curve Ott (2014).

2.7 Conclusions

In this section we analyzed the strategic behavior of firms competing in electricity capacity auctions. We began with a simple case with no competition and expanded to multiple bidders. In the no-competition situation with only a single time period, bidders were honest about their costs. However, once in a multi-period horizon, a bidder who had won in the first period would bid 0 in every subsequent period. Such bidding makes it more difficult for new capacity to enter as it pushes the installed capacity higher and thus the capacity price downward. Under imperfect information, participants have an incentive to underbid in attempts to win the auction.

With multiple bidders, we were able to find closed form symmetric equilibrium optimal bid functions for each bidder. Under the private information and common prior assumptions, we found that participants would shave their bids to beat out the other participants. The severity of this underbidding depended on the payoff structure, with the demand curve auction having the closest to truthful bids. The uncertainty of the payment is what drives the untruthful bidding. If it is known, then respondents truthfully report their costs.

From the perspective of the auction holder, more participants led to greater additions of capacity in all permutations of the auction. Because the capacity price comes from the ISOs own demand curve, the prices settled in the auction are “fair” to the auction holder. Nonetheless, the auction holder overpays when the supply curve is not smooth. Compared to the flat price auction, the demand curve auction results in lower expected gains to the winners, and lower overall procurement of capacity.

In the simulated scenario comparing the two auction types, fixed winners and decreasing demand curve, we see that at higher VOLLs, the flat demand curve results in lower expected costs across all possible prices compared to the sloped demand curve. At low prices and very high prices the fixed winner auction results in lower costs. Were the VOLL to grow ever higher, the sloped curve would eventually result in lower costs compared to the flat case at high prices. Conversely, at low prices, the sloped case would result in astronomically high costs to the system.

Chapter 3

Simulation

3.1 Introduction

Now that a theoretical analysis of a demand curve based capacity auction has been discussed, this chapter analyzes this behavior on a larger scale, that is, with more participants and over an extended period of time. Contrast this with the work in chapter two that only dealt with a few players and was a single period game. An agent-based simulation model of electricity markets will be used to model such an auction as this continues the bottom-up style approach of the second chapter, where each generator acts independently of others in maximizing its objective. Here the generator's objective is to maximize their revenues from operating in electricity markets. The model framework is developed on top of an existing open-source model called AMES (Agent-Based Modeling of Electricity Systems) developed out of Iowa State University. The AMES model simulates an electricity market with a day-ahead electricity market where economic conditions-based market entry is permitted, as well as a demand-curve based capacity market. For more detailed information on the workings of the AMES model, see Sun and Tesfatsion (2007). A quick introduction to the energy system and mechanics will follow below.

3.2 Transmission Grid

The transmission grid itself consists of at least two nodes and at least one branch. A node can have a multiple generation sources, a demand sink, or both. The nodes must all be connected by branches, but a node does not need to be connected to every other node by a branch. Each node is connected in the network such that power can flow from one node to any other node, that is, there are no islands. See Figure 1 for an example of various grids allowed in the model.

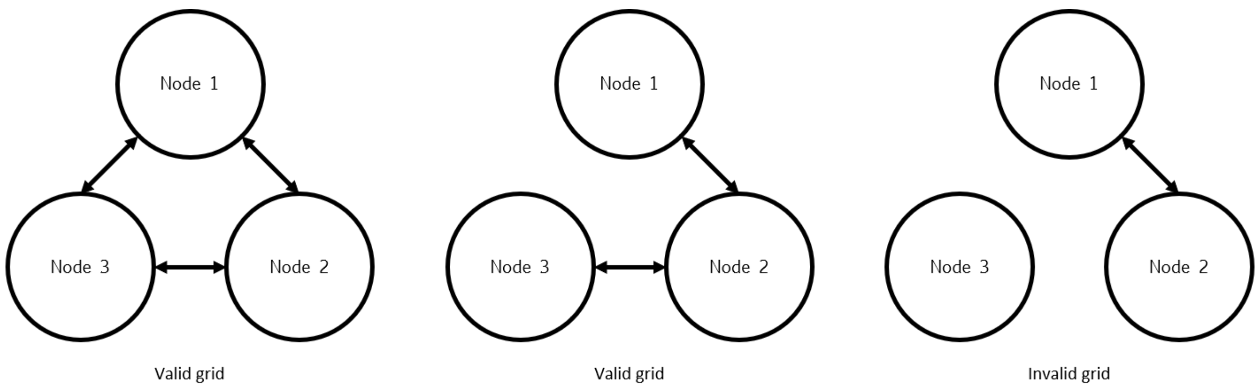


Figure 3.1: Example grids in the AMES model

The branches have a capacity constraint that limits how much power can be transmitted, leading to situations where certain branches cannot transmit any more power, affecting prices and generation. Cheaper generation that could have been dispatched may be bottlenecked due to transmission constraints, forcing more expensive generation from a source that is not bottlenecked to produce power.

3.3 Generators

The generators are also located at nodes in the grid, and it is possible to have multiple generators at the same node. Each generator i has a minimum and maximum hourly power production capacity. For each generator, hourly production level r must be between

$$Cap_i^L \leq r \leq Cap_i^U$$

Each generator has a total cost function that gives the total cost of production per hour (\$/hr) for an hourly production level r. a_i and b_i are constants, and c_i is the fixed cost for generator i.

$$TC_i(r) = a_i \cdot r + b_i \cdot r^2 + c_i$$

The constants are given exogenously by the user. The marginal cost is then given by

$$MC_i(r) = a_i + 2b_i \cdot r$$

At each day D the generators submit their marginal cost function over a production interval (lower limit and upper limit of production) for each hour of day $D + 1$. The offers the generators submit can be different from the true production limits and marginal cost curve. If the agents are allowed in the model to learn strategic behavior, then they can report untruthful marginal costs and production limits. For now we consider the no-learning case. Only truthful reporting of costs is permitted. The price of electricity in the system is determined by the marginal cost required to produce the last unit of electricity required to meet demand. Because of transmission constraints, it is possible that the lowest cost generation available cannot be used to meet supply at all nodes. Thus the Locational Marginal Prices (LMPs) were given their name. LMPs can be different at different nodes. If there were no transmission constraints, then there would be a single electricity price in the system.

3.4 LSEs

The load serving entities are located at different nodes in the grid (not every node has an LSE, and a node doesn't have more than one LSE) and submit a daily load profile each day D for day $D + 1$. The profile consists of the demand for electricity that each LSE must serve for each of the 24 successive hours in day D+1. The demand can be either fixed, or have a price-sensitive component enabled.

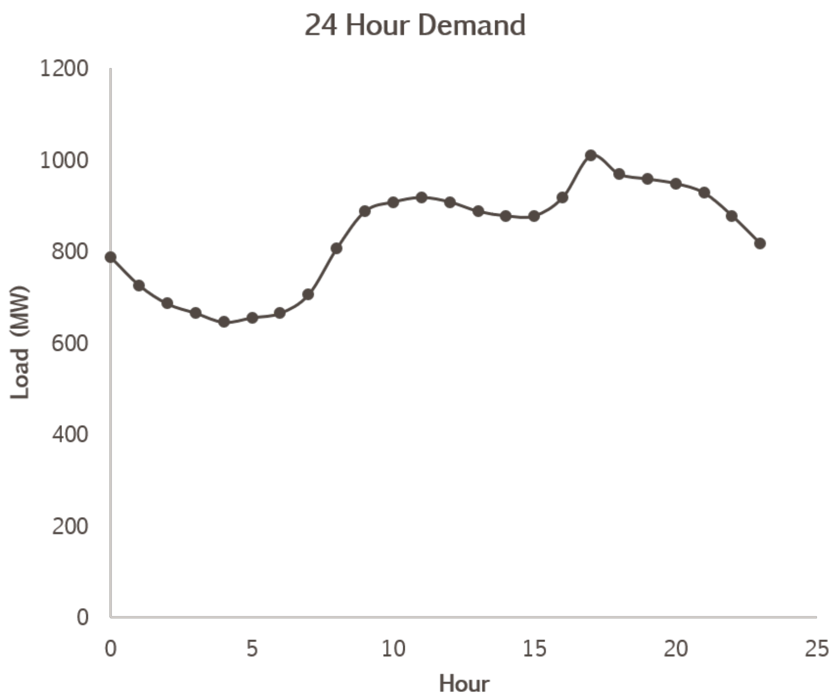


Figure 3.2: Example load shape in the AMES model

Figure 3.2 shows a plot of demand profiles for an LSE. The demand is lowest in the middle of the night, between 2 am and 6 am, and peaks at 5 pm as people return home and begin to use appliances, TVs, water heaters, etc. In the original AMES model, load is fixed, and the generation resources are sufficient to meet the load. To test the efficacy of capacity markets and energy-only markets, the load must increase to put pressure on reserves and generation resources. In our modification of the AMES model, demand for each hour is assumed to grow at rate τ per day.

3.5 Unit Commitment

The Independent System Operator (ISO) must determine for each day $D+1$ a schedule of power production and prices from the information on day D. Many factors go into the creation of this schedule, including generator supply curves, load serving entity demand bids, limits on transmission branches, and the balance of supply and demand at each node. The ISO then must solve an optimization problem that results in the electricity demand being met for the lowest possible total cost to the system.

The optimization problem is one that is common in actual wholesale electricity markets. While the market itself is an alternating current (AC) optimal power flow problem (OPF), in practice this is quite difficult to solve. Thus it is common to instead use a direct current (DC) approximation to the AC power flow optimization problem. In the DC OPF, certain constraints are simplified allowing other constraints to be relaxed or removed. This makes for a problem that can be solved more efficiently.

The optimization's objective function is to minimize the total production cost of all generators throughout the system while meeting electricity demand at all nodes. The constraints in the problem are the generator production constraints, both minimum and maximum production levels, and the supply-demand power balance constraint. The optimization problem is a convex quadratic program, for which there are many available solvers.

Sun and Tesfatsion make a small modification to the objective function by allowing for a penalty term on the squared voltage angle differences. See Sun and Tesfatsion (2007) for the full details of their algorithm for solving the DC optimal power flow problem; the algorithm is unchanged in this modified version of AMES.

3.6 New Entrants

As demand in the electricity market grows, the existing generation capacity will be insufficient to serve load. As more expensive generation gets dispatched to meet load, prices rise. A prospective entrant to the market would observe such high price signals, and, should prices be high enough, decide to build a new power plant in the market. The market price would need to remain high enough after entry to allow the newly constructed plant to recover its capital costs.

One way to measure the competitiveness of such plants is with the levelized cost of electricity (LCOE). This is a dollar per megawatt-hour cost of constructing and operating a power plant over its serviceable lifetime. Different generation technologies have different levelized costs, but those with the same primary movers (i.e. steam generator, gas turbine, internal combustion, combined cycle, etc.) and fuel sources are similar. The inputs to LCOE are capital costs, fuel costs, operation and maintenance costs (both fixed and variable), and the generator's utilization rate. Renewable technologies tend to have lower or negligible fuel costs. Nuclear power has low fuel costs but extremely high capital costs, in comparison to gas-fired combustion turbines which have a low capital costs and a fuel cost that in recent years has been quite low as well.

The computation of levelized cost on a yearly basis can be simplified as follows:

$$LCOE = \frac{CapCost + FOM}{ExpectedGeneration} + VOM + Fuel \tag{3.1}$$

For the purposes of this paper we ignore the discount rate. FOM and VOM are the fixed and variable operations and maintenance costs, respectively. The expected generation is an estimate of megawatt-hours that a particular plant will run during the year. The capital cost described here will be defined as the overnight cost (the cost of building the plant in a single moment, were it possible) divided by the operating lifetime of the plant.

The LCOE can be considered the average cost of a megawatt-hour of energy generated by a power plant. We can turn our efforts toward conditions for market entry and attempt to determine when expected future revenues exceed expected future costs. Recently, the Energy Information Administration has developed a metric for comparing the cost of a power plant and the revenues of that power plant (Namovicz (2013)). This metric is called the levelized avoided cost of energy (LACE). The LACE of a plant can be considered the potential revenue per MW. It represents the amount and cost of generation displaced. To that end, LACE can be computed as

$$LACE = \frac{\sum_t p_t \cdot dispatch_t + CapacityPayment}{ExpectedGeneration}. \tag{3.2}$$

The expected generation in the denominator is defined as seen before in the LCOE formula. The numerator represents the price of electricity p times the dispatched generation in hour t for an entire year. The capacity payment is added as an additional source of revenue. It is important to recognize that the capacity payment is not guaranteed until the generator submits a bid into the capacity auction and clears in the auction. Thus the generator won't receive any capacity payments for the current calendar year when entering at any time other than during the capacity auction itself.

If the LACE is greater than the LCOE, then it is a net positive to construct the plant and enter the market, since its revenues outweigh its costs. Each new power plant that enters the market lowers the LACE of a subsequent plant, since the first plant displaces more expensive generation. This will continue until the LCOE of any new plant is higher than the LACE, and new entry stops. The LACE can increase as demand grows or existing plants retire, and again new plants will enter until equilibrium is reached. A competitive market will have LACE=LCOE holding everything constant.

As discussed earlier, dispatch and marginal prices are calculated using the DC approximation of optimal power flow. For our LACE and LCOE calculations, we must calculate an estimate of generation and marginal prices. Unfortunately, running the quadratic program to compute day ahead hourly dispatch for a year in advance would be computationally expensive. Therefore, a simplified approach was developed based upon the idea of merit order dispatch.

In the model there are three generation technologies: coal, combined cycle, and combustion turbines. The marginal cost for each plant of a given technology is the same. Furthermore, the marginal cost as a function of fractional generator output is non-overlapping for each technology. That is to say, all coal will be dispatched before any combined cycle, and all combined cycle will be dispatched before any combustion turbine generation. Because of this, an extremely simple merit order dispatch model can be applied. The algorithm is as follows:

Sum installed capacity (IC) for each generation type making sure to include the prospective plant's capacity, IC_{coal}, IC_{CC}, and IC_{CT}. For a given load level, subtract from it the cheapest generation source, in this scenario, coal. If the coal is sufficient to meet all demand, the marginal cost can be calculated from the fraction of coal dispatch required over installed coal capacity. If the coal was not sufficient the leftover load is distributed amongst the combined cycle resource, and marginal price calculated. This repeats until all generation types are exhausted. An example will make this more clear.

A generation technology's variable fuel cost can be represented as

$$VFC = Aq^2 + Bq + C \tag{3.3}$$

where q is the fraction of power output as opposed to output in MW. A 10 MW plant is the same as ten 1 MW plants. Generating 8 MW with a 10 MW plant costs as much as 0.8 MW each from ten 1 MW generators.

Imagine there are 200 MW of coal generation available, 350 MW of combined cycle generation available, and 50 MW of combustion turbine generation available. A prospective plant hopes to add 50 MW of combined cycle generation. If the total load is 400 MW, then 200 MW of coal are dispatched and 200 MW of combined cycle are dispatched. 200 MW of CC generation is 50% of the installed generation (including the prospective plant). Because every generator of a particular technology has the same marginal cost in terms of fractional output, then each plant that makes up the 400 MW will be dispatched at 50%. The marginal cost would be $2A(.5) + B$. This procedure as a whole is a fast computation and is performed for each of the next 8760 hours to obtain the expected generation and LMP inputs to both LACE and LCOE.

The predicted capacity price is based on the ISO demand curve for capacity, with installed capacity set at the current level plus the prospective generator's added capacity, and the peak demand set to be the next year's peak. An illustration of the process with PJM's demand curve can be seen in Figure 3.

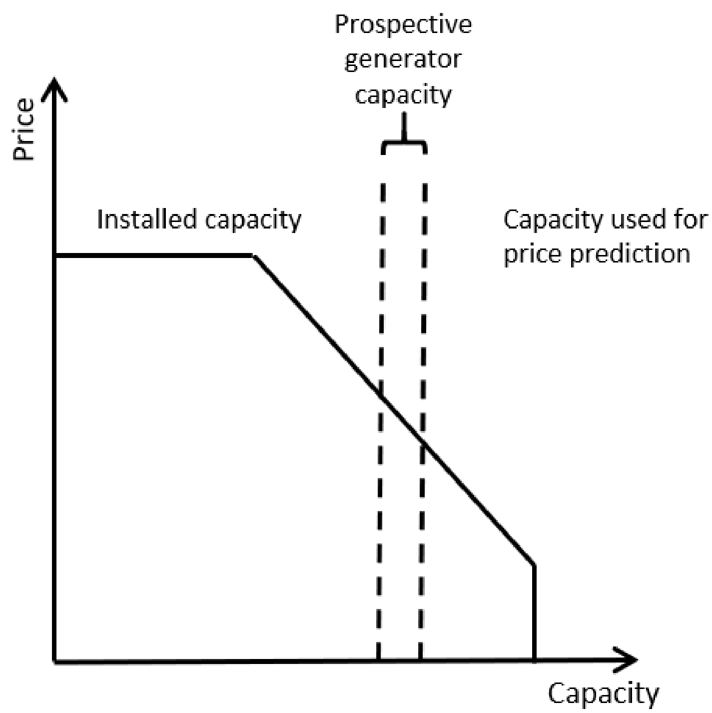


Figure 3.3: Predicting capacity price.

A capacity market is not the only way that generators can enter the auction. If market conditions indicate that it would be profitable to enter the market, then the model allows for a generator to enter. Of course, entrants based on market conditions can still participate in the capacity market as an existing generator. They did not need the guarantee of a capacity payment to provide the necessary subsidy to enter the market. Once the generator enters, it participates in the day-ahead market like any other existing generator, and is free to enter the subsequent calendar year's capacity auction.

3.7 Capacity Market Overview

Every year, a capacity auction is held by the ISO. In such an auction, every year all existing and potential new generators bid a price-capacity pair. A supply curve is constructed from these bids and intersected with a demand curve for capacity set by the ISO. This demand curve is piecewise linear and decreasing. The kinks in the demand curve occur at reserve margin targets set by the ISO, such that capacity levels below the target margin have a very high capacity price (almost double the cost of new entry), while capacity levels that exceed the target reserve margin by a large amount have a price of 0.

Those clearing the auction receive the capacity price for that year for each unit of capacity supply they entered into the auction. All generators currently active in the system must participate. At the moment in the model, existing generation bids in a capacity supply offer of \$0/MWh. Historically, it has been observed in PJM that existing generation bids in their capacity at a price of 0, so this is not an unreasonable assumption. Within this model, the demand function is easily modified and multiple shapes of curves can be compared to determine their efficacy.

In addition to the existing generation participation, there are new prospective generators that would agree to be built if the capacity clearing price is high enough. The number of new potential entrants is set to a fixed number. The costs and size of the generator are normal random variables. The new generation bids in their cost of new entry minus the ISO-set expected revenues. They are, in effect, bidding their true net cost of new entry.

When the auction clears, if one or more of the new entrants cleared, they come online that year. They then behave like any other generator.

Every day, those that cleared in last year's auction receive a capacity payment. Since unforced outages are not implemented in the model there is no check for whether or not a generator is able to meet its obligation to the capacity market; it is assumed to do so.

3.8 Texas Energy Market

The electricity market in Texas is managed primarily by the Electric Reliability Council of Texas (ERCOT). ERCOT serves about 90% of the load in Texas, and its grid covers about 75% of land area in Texas. There are over 570 generation units in ERCOT, and in 2016, generation capacity was 52% natural gas, 22% coal, 20% wind, 6% nuclear, and 1% solar, hydroelectric, biomass, and other sources.

We make a number of simplifications of the ERCOT grid to ease computational requirements. Generator data from ERCOT and the Energy Information Administration were analyzed and aggregated into three main zones, represented by nodes in our model, with three generators in each zone, representing coal, combined cycle, and combustion turbine generation sources. The generator sizes are proportional to the total generation of that type in that zone.

A rough estimate of generation capacity in megawatts (MW) by type by zone is given in Table 3.1.

In the model the three types of generators with their sizes in MW appear at the three nodes representing Houston, North, and South as shown in Table 3.2. The sizes have

	Coal	CC	CT
Houston	2504	7948	1987
North	10589	13724	3431
South	5466	9543	2386

Table 3.1: Generation capacity by type in select ERCOT zones

been scaled so that in the model they can each be represented by one generator, to save on computational load. This gives an initial installed capacity of 5750.

	Coal	CC	CT
Houston	250	800	200
North	1050	1370	340
South	550	950	240

Table 3.2: Scaled capacity by type in select ERCOT zones

Seasonal demand was calculated by averaging maximum daily load relative to January 1st, for the years 2011-2016. Electricity demand usually has a sinusoidal shape-demand and is higher in winter months due to heating and summer months due to air conditioning. Because Texas has much hotter summers and milder winters, the summer peak is almost 1.5 times as high as the winter peak. The maximum load on January 1st, 2016 was 41171 MW, for reference. Electricity demand was fit to the model

$$\frac{Demand_t}{Demand_{Jan1}} = a\frac{\sin\left(\frac{\pi(t-b)}{w}\right)}{\frac{\pi(t-b)}{w}} + c \tag{3.4}$$

In Figure 3.4 we see the fitted curve overlaid on top of the averaged historical data.

Variable	Estimate
a	-7.51
b	-824.76
w	138.50
c	1.29

Table 3.3: Coefficient estimates

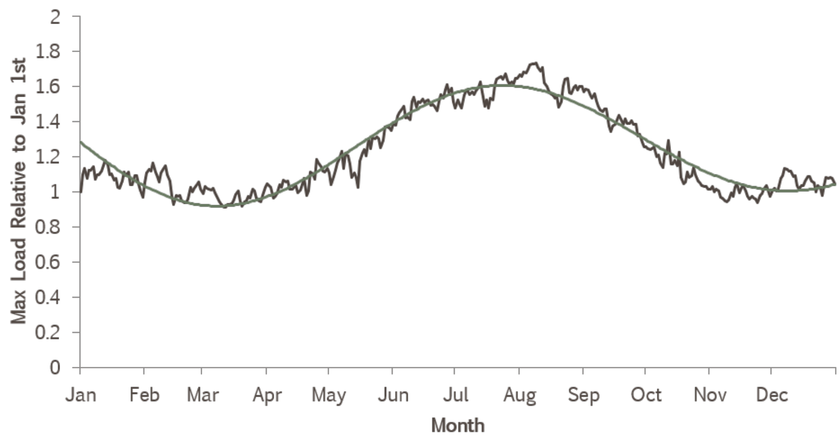


Figure 3.4: Historical versus modeled load relative to January 1st

The parameters were adjusted so that the first and last day of the year had similar levels of electricity demand. The curve is shifted by +19 days so that the 1st and last days of the year are more similar, better representing the cyclical temperatures. This function peaks at about 1.6 times the value of the initial load on the first day of the year. It should be noted that the ends do not match up exactly, and so there are small jumps every year from December 31st to January 1st of the next year.

Moving to daily granularity, the hourly load peaks at hour 17 as seen in the last section, and the actual load at each bus is as follows in Table 3.4:

	Hour 0	Hour 1	Hour 2	Hour 3	Hour 4	Hour 5	Hour 6	Hour 7
Houston	787.58	726.67	686.41	666.11	646.17	656.15	666.11	706.73
North	1837.69	1695.56	1601.63	1554.27	1507.75	1531.01	1554.27	1649.04
	Hour 8	Hour 9	Hour 10	Hour 11	Hour 12	Hour 13	Hour 14	Hour 15
Houston	807.52	888.39	908.69	918.66	908.69	888.39	878.42	878.42
North	1884.21	2072.92	2120.28	2143.54	2120.28	2072.92	2049.66	2049.66
	Hour 16	Hour 17	Hour 18	Hour 19	Hour 20	Hour 21	Hour 22	Hour 23
Houston	918.66	1009.50	969.24	958.91	948.95	928.65	878.42	817.87
North	2143.54	2355.50	2261.57	2237.47	2214.21	2166.85	2049.66	1908.36

Table 3.4: 24 hour load data for the model

	Coal	CC.	CT
A	0.6697	9.7259	2.259
B	17.115	19.094	41.538
C	1.3555	11.015	9.9821

Table 3.5: Variable fuel cost parameters in the model

Though there does exist load in ERCOT's South zone, we choose not to have any here for modeling purposes. The reasoning is two-fold. One, to simplify the model for computing purposes. Two, in the case of transmission constraints, it would be possible with that South zone generation capacity goes under-utilized without a demand sink present at that bus.

At hour 17, the peak demand is around 3300 MW on the first day of the model. This means that in the first year, the maximum demand will be around $3300 \times 1.6 = 5280$. The installed reserve margin is then about 9% (5750 MW/5280 MW). Again, using historical data, demand growth in ERCOT has been about 1.1% per year, so this reserve margin will shrink as time goes on unless new supply is added to the market. The transmission grid connects all three nodes pairwise, with a maximum capacity of 9999 MW per line. A diagram of the initial grid can be seen in Figure 5.

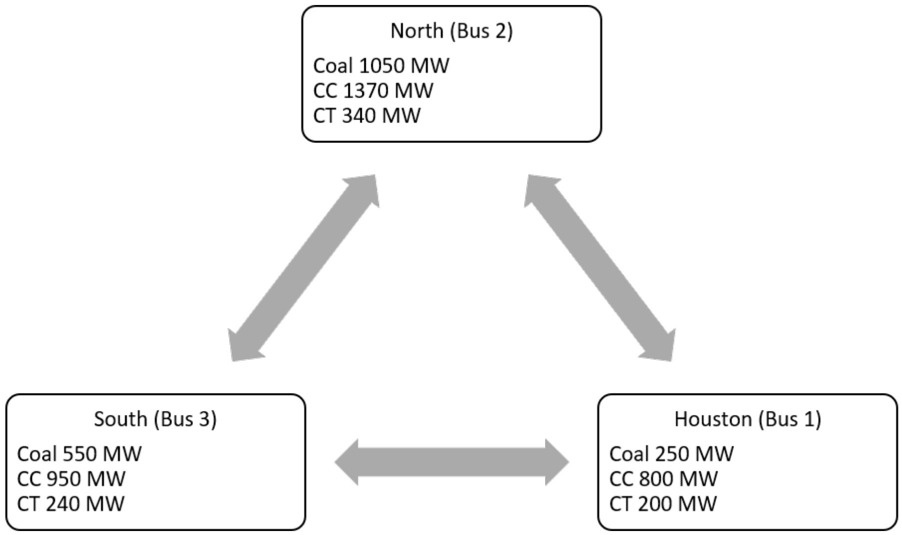


Figure 3.5: Overview of generation in the model

A	9.7259	a	0.014
B	19.094	b	19.094
C	11.015	c	7159.75

Table 3.6: Parameters for a 650 MW combined cycle plant

Generators have the variable fuel cost curve parameters shown in Table 3.5 where $VFC_i(q) = A_i q^2 + B_i q + C_i$ is given in \$/h/MW capacity and q is the fraction of generator output. Recall that the model uses cost curves of the form

$$TC_i(r) = a_i \cdot r^2 + b_i \cdot r + c_i \tag{3.5}$$

where total cost is given in \$/hr. To convert this variable fuel cost to a fractional output form the model can use, we have to transform the parameters A, B and C. Noting that $q = r/Cap$, we can set $a_i = \frac{A}{Cap_i}; b_i = B_i; c_i = C_i \cdot Cap_i$ For example, for a 650 MW combined cycle plant, Table 3.6 shows how the parameters translate.

The average cost per hour per megawatt at different fractions of output is shown in

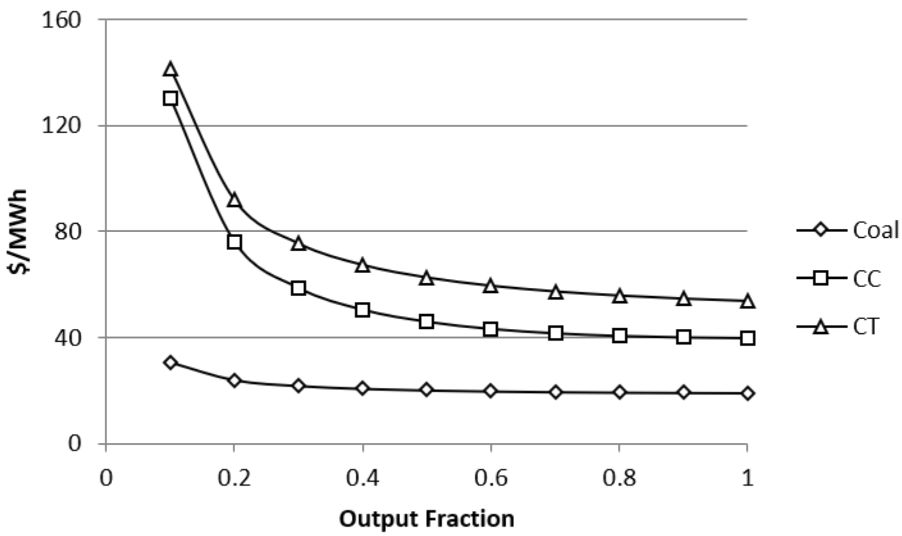


Figure 3.6: Average cost per MWh as a function of output fraction for three generation technologies.

Figure 3.6, obtained by multiplying the average heat rate by the fuel price for each technology.

Note that the plants become more efficient the closer they operate to full output. The combined cycle and combustion turbines have very high startup costs, as can be seen by the dominating effect it has at low levels of output. Coal is relatively cheap compared to the two natural gas fired plants. In the current model, the coal price is \$2.18/mmbtu and the natural gas price is \$6.29/mmbtu. The two natural gas technologies are very costly at low levels of output, and while they do not overtake coal, they do become more competitive.

The incremental cost curves or marginal cost curves for the three technologies can be seen in Figure 3.7. It is easy to see how the fast dispatch described in the previous section is accomplished. Since none of the marginal cost curves overlap, one can determine the marginal price as if reading the below graph from bottom to top. First, coal is dispatched

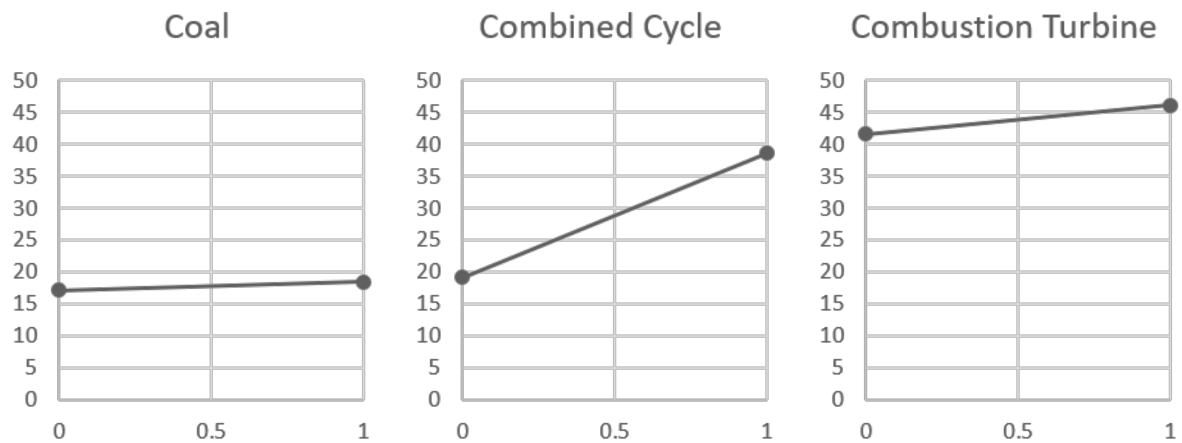


Figure 3.7: Marginal costs for three generation technologies in the model.

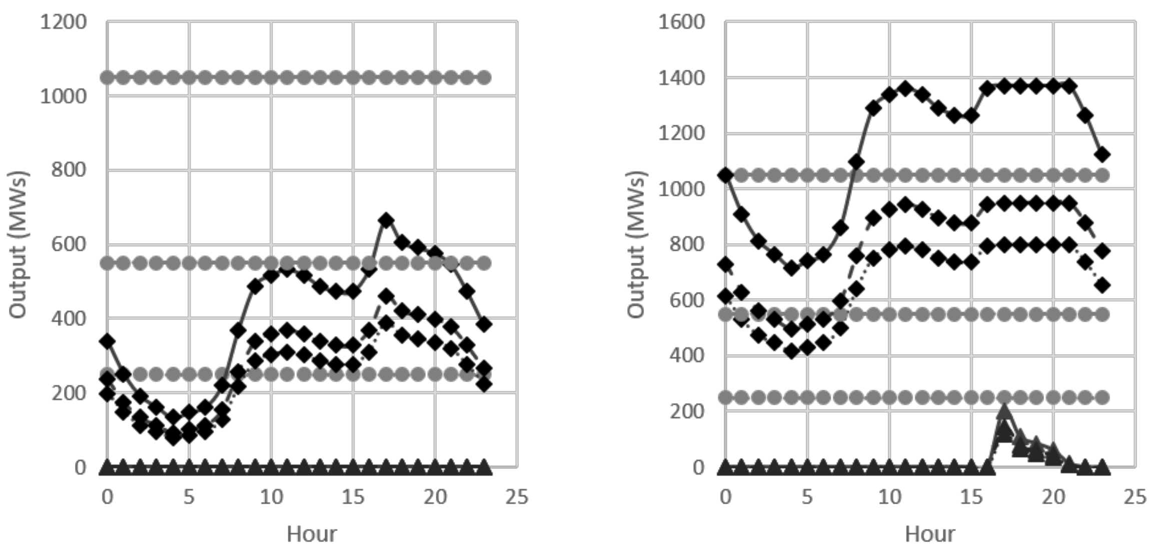
up to 100%, then we move to the right as we dispatch CC, and then finally CT. In the absence of transmission constraints, it is always cheapest to dispatch coal before any combined cycle, and any combined cycle before combustion turbine.

It can be noted from this figure and the one preceding it that the marginal prices that come out of the model will cap out around $45/MWh. Any combined cycle plant operating at under 50% output, and a combustion turbine plant at any level of output will be operating at a loss. Note also that this is at peak prices, meaning the economics are much worse when demand is at lower than peak levels. We will see the impact of these curves in the scenarios to follow.

3.9 Base Case

In the base case, the demand grows at a rate of 1.1% per year, or about 0.03% daily. There are no transmission constraints, so we will not expect to see any differentiation in the locational marginal prices at the three nodes. New entrants are not permitted to enter. The price for electricity when demand cannot be met is set to be \$47, just higher than the marginal cost of the most expensive generation available (the combustion turbine).

Figure 3.8a illustrates an initial benchmark run of the first day shows the typical load profile with generation peaking at hour 17. Coal is represented by dots, combined cycle by squares, and combustion turbine by triangles. As expected, coal is dispatched steadily, as it is the cheapest resource available. The combined cycle generation rises and falls to meet the demand as necessary. The combustion turbines are never dispatched, as the demand is not high enough in January when the model is initialized.



(a) Commitments on day 0

(b) Commitments on day 188

Figure 3.8: Generator commitments on day 0 and day 188. Coal is represented by circles, combined cycle by squares, and combustion turbine by triangles.

Contrasted with the Figure 3.8b on the right, which is the generation commitment

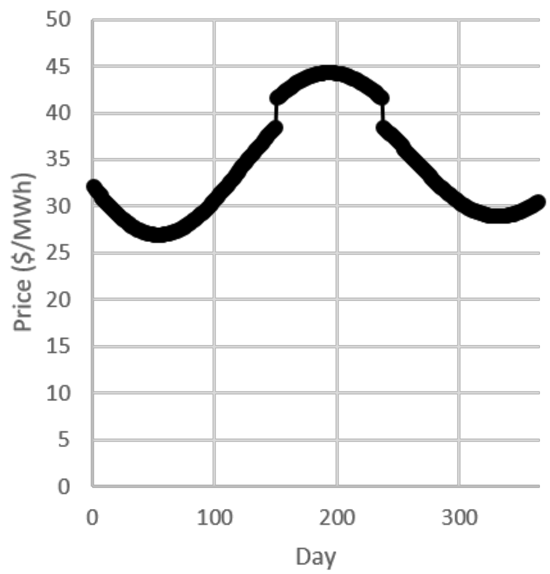


Figure 3.9: Locational marginal price at hour 17

schedule for the highest load of the year (day 188, meant to approximate the hottest days of the summer). At hours 17 through 21, even the combined cycle is at its maximum, requiring the combustion turbine peaker units to come online for four hours. We can see the location marginal price at hour 17 for the year reflect this, shown in Figure 3.9. Recall that in the marginal cost curves earlier there was a gap between the CC and the CT curves, a gap that is reflected in the graph of the LMPs for hour 17 for the year.

In Figure 3.10, the profits for one year can be seen. Recall that in the average costs diagram natural gas fired plants would operate at a loss unless the price was very high. We do see that the profit for combined cycle (bottom 3 lines) per day is higher during the summer, but even though it may be positive during peak hours (hours 17-21), the other hours of the day still net a negative return for the day as a whole. The combustion turbines

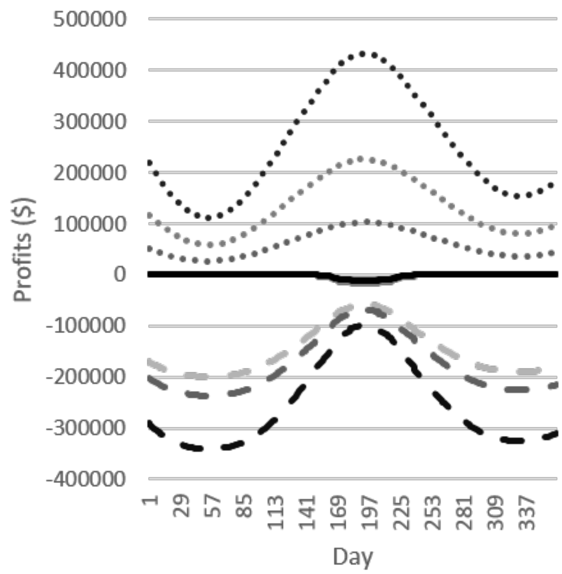


Figure 3.10: Generator profit by day. Dotted lines represent coal, dashed CC, and solid CT.

(middle three lines) lose money when they come online as even though their marginal costs are being cover, their fixed costs are not. The coal plants make a profit every day (top 3 lines).

Beginning in year 7, the 1.1% per year demand growth causes peak load to be higher than installed capacity. There are 24 straight days of outages at hour 17 in year 7. A total of 168 hours did not meet electricity demand, which was, in total, 17.7 GWh of unsatisfied demand.

The total cost of supplying as much energy as possible was \$9.93B. At a current value of lost load in ERCOT of \$9,000 per MWh (Surendan et al. (2016)), the impact of the loss was \$158M. Should a more pessimistic view be taken, and assume that the loss results in a system-wide blackout for that hour, the total served electricity costs would be \$9.89B and the value of the lost load \$8.85B. The reserve margins are expectedly poor, starting at 6% and decreasing steadily by a little over 1% per year.

3.10 Energy Only Entrants

To incentivize new generators to enter, we change the maximum price paid from \$47/MWh to various levels, including \$1,000, \$5,000, \$10,000, and \$20,000 when load cannot be met. At a value of \$1000 no generators enter the market, resulting in a situation identical to the base case, with the exception that more money is paid (since the cap is now \$1000 vs \$47), resulting in a higher overall system cost. In all, this is not a worthwhile change to be made to the system.

For the remaining cases, the net value to a generator ($LACE - LCOE$) doesn't become positive until very late in the simulation. In the \$5,000, \$10,000, and \$20,000 cap scenarios, the first generators are built before any loss of load event. Table 3.9 at the end of this section with the key metrics for the different scenarios shall shed some light. For now, we will discuss the \$20,000 price cap scenario.

In this scenario, 10 new combined cycle generators are built. There are 2 hours where demand cannot be met, accompanied by a loss of just under 4 MWh of demand. Compared to the 17.7 GWh from the base case, consumers and the system operator should be much more satisfied. In addition, while the cost to serve the load is higher, it is not by much, a difference of \$131 million.

Instead, should we consider the pessimistic scenario where all load is lost in that hour, the total impact of the loss is only \$131M, with a cost of serving load of \$9.83B for the

Year	Reserve Margin	Installed Capacity	Peak Demand
1	5.7%	5750	5440
2	4.6%	5750	5500
3	3.4%	5750	5559
4	2.3%	5750	5618
5	1.3%	5750	5677
6	0.2%	5782	5737
7	0.5%	5851	5796
8	0.4%	5905	5855
9	0.2%	5955	5914
10	0.4%	6021	5973
11	0.5%	6061	6033
12	0.4%	6139	6092
13	0.1%	6187	6151
14	0.0%	6246	6210
15	0.1%	6301	6270

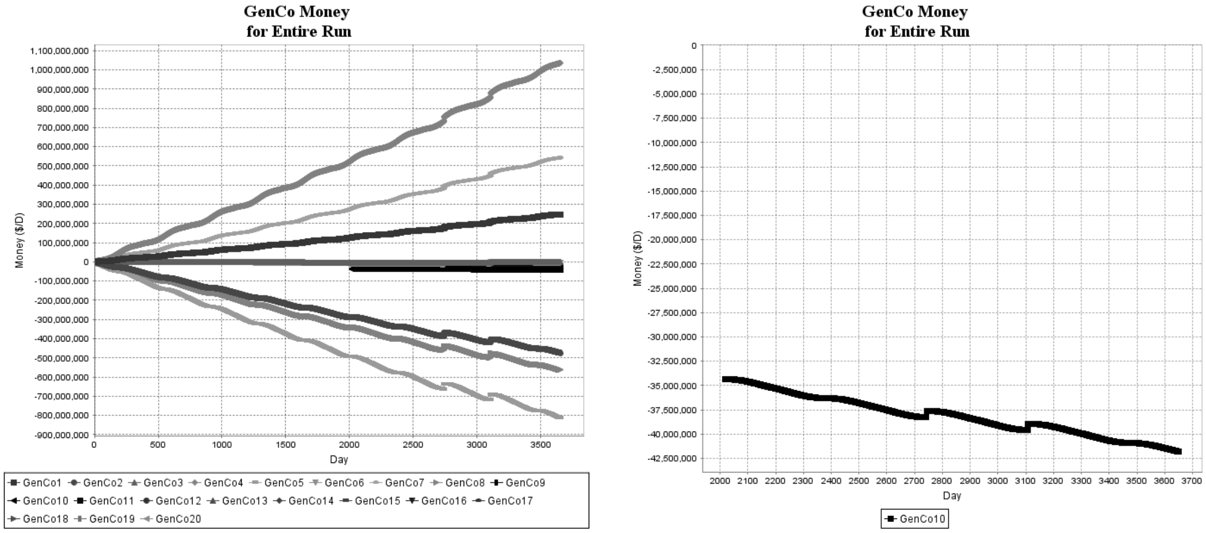
Table 3.7: Reserve margin for a 15-year run, $20,000/MWh price cap.

remaining hours. This is not surprising since there are 2 hours of lost load in this scenario compared to the 168 hours lost in the base scenario. It should also be noted that this price cap achieves a less than 1 day in 10 year loss of load expectation, commonly held as a benchmark by many system operators (citation).

The reserve margins for a 15-year run at the $20,000/MWh cap is shown in Table 3.7. The system enters a steady-state, reserve margin holds steady at around the 0.5% mark, varying slightly due to the randomly drawn capacity of new generators.

Note that in Table 3.7, the reserve margin is not actually less than 0%. While enough capacity is built to handle the peak demand in the year, the outages occur a few days before new generation is built for those peak days. One may wonder why, at a price cap of $20,000, there are still outages? In fact, the price cap could be set arbitrarily high, and some load would still be lost. This is because for the price to ever actually hit this cap, the demand must be greater than the supply. If this is not the case, then generators will not actually ever receive the massively inflated prices and would not have chosen to enter the market. Because we assume that generators can perfectly anticipate demand and act accordingly, prices will always be anticipated to hit the cap for there to be new entry in an energy only market.

The cash holdings of each generator may help illuminate this point. All the natural gas technologies lose money. The 6 starting generators have no say in the matter; a prospective generator knows this and will not enter as doing so would only guarantee a loss. If, however, every summer the prices skyrocket, those brief periods of operation will allow the generator to recover some of those costs. In the graph seen in Figure 3.11a, we see that the coal plants have no problem generating a profit (the three topmost lines), as their total costs of operating are lower than all LMPs throughout the simulation period. All other generators incur losses as a result of prices not rising high enough. You can see two small jumps in Figure 3.11a and Figure 3.11b after day 2500 where the outages occurred and prices hit $20,000/MWh.



(a) Generator cash holdings for entire run (b) Generator 10 cash holdings since entry

Figure 3.11: Generator cash holdings

Focusing on a single new entrant, the first generator that enters the market does so on day 2018, at a total cost of $34M. By day 3650, they are at holdings of -$41M, a net loss of -$7M. It is clear from this figure that there is no hope of ever earning a positive return on this investment.

If this is the case, then why did so many generators enter? A total of 10 in this $20,000 cap case. The reasoning is that when a generator is considering entering, it looks one year ahead and computes the operating and capital costs it must pay and subtracts that from its projected earnings. In all these situations, the generator foresees the outage and the high price being paid out, which more than covers the rest of the year. The price caps are orders of magnitude greater than the average LMP the rest of the year. The problem is that once this generator enters and is waiting for excess demand and the price cap, another prospective generator is doing the same calculation and makes the same decision to enter, predicting shed load sometime in the next year. So, it enters and unfortunately, this happens before the first generator ever experiences prices hitting the cap. With the addition of two generators in that original year time period, demand does not outpace supply. This happens for every generator that enters, with the exception of two hours in the 10-year period, which as we saw in the cash holdings graphs, is not enough to make a significant impact.

A large part of this is that the generators entering have capacities drawn from a normal distribution with a mean of 30 MW. At a growth rate of 1.1% per year, and peak demand in the mid-5000 MW range, that equates to a growth in peak demand of 60 MW, requiring about two generators to enter, which is just enough when staggered to cause the leapfrogging effect we see. At the end of this chapter we will investigate a different new generator default size to gauge what impacts it has on the system.

3.11 Capacity Market

With the capacity market active, there is not a single day of load shed. The LMPs at hour 17 seen in Figure 3.12 get high but not quite as high in the other two cases, and never hit the price caps. The small caps are the points at which the combustion turbines were dispatched. Note that they shrink as more combined cycle is built.

The market initially starts at a 5% reserve margin, and the ISO in our example has a target reserve margin of 0%. The target of 0% is set so that the capacity market case is similar to the energy only market cases. In reality, ISOs have a target well above 0%, to account for unforced outages. Because we do not have any supply shocks in our cases, we treat the target reserve margin of 0% as sufficient for lossless generation. We see in this model that the capacity market does its job in raising the reserve margins to be at or slightly

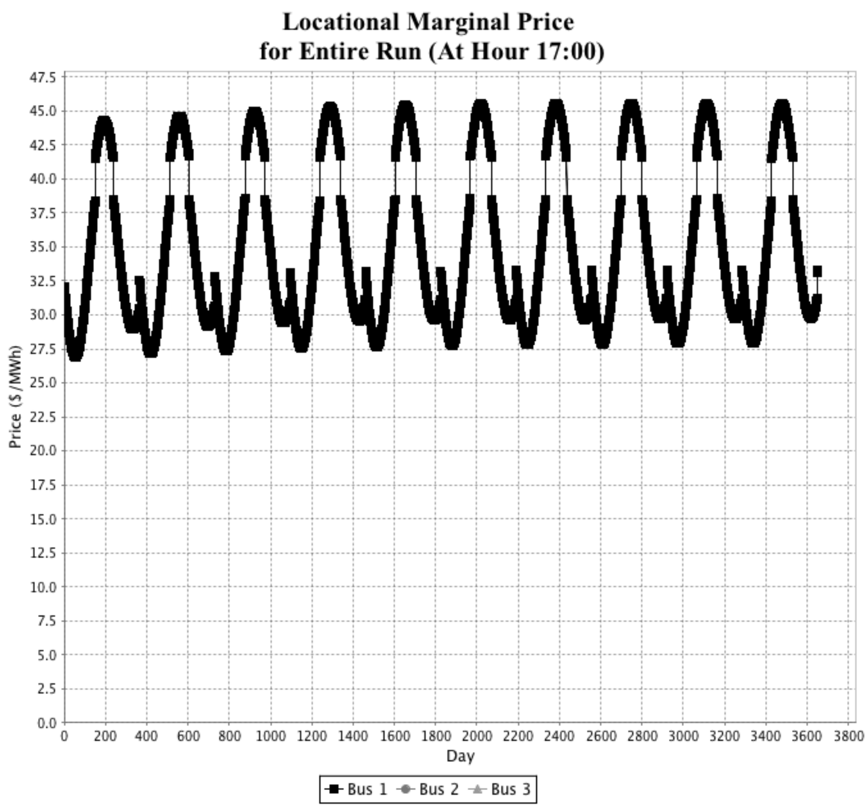


Figure 3.12: LMPs for hour 17 over entire run

Year	Reserve Margin	Installed Capacity	Peak Demand
1	5.6%	5750	5446
2	4.5%	5750	5505
3	3.3%	5750	5564
4	2.2%	5750	5624
5	1.2%	5750	5683
6	0.7%	5780	5742
7	0.9%	5855	5801
8	1.0%	5917	5861
9	0.5%	5952	5920
10	0.5%	6014	5979

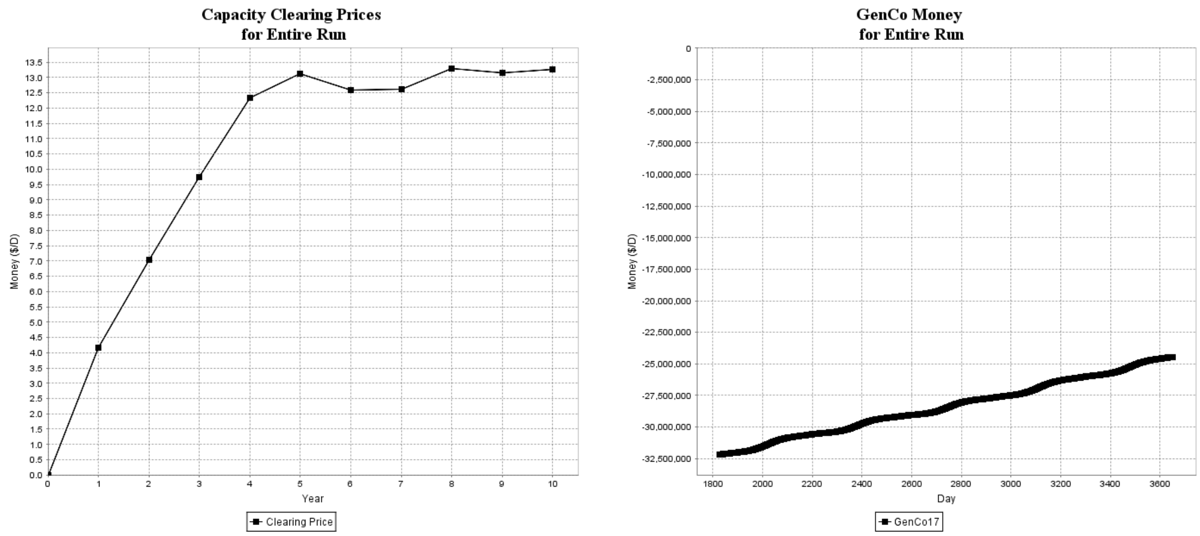
Table 3.8: Reserve margins in capacity market case

above the 0% level; the results can be seen in Table 4.

The interesting thing to note is that because the LACE considers potential capacity revenues, no generators came online in the capacity auction. The results are quite stable and would repeat indefinitely. Small variations in generator size could potentially cause dips but nothing that would disrupt or destabilize the system. The main purpose of the capacity market is to bridge the gap between the revenues gained in the energy market and the total costs associated with new entry such as fixed costs and capital costs. In this model, we set the net CONE to be the difference between the LCOE and the LACE, adjusted by the capacity factor.

$$NetCONE = (LCOE - LACE) \cdot \frac{ExpectedGeneration}{Capacity \cdot 8760} \tag{3.6}$$

We see in Figure 3.13a on the left that the price is relatively steady around \$13.5/MWh. On the right in Figure 3.13 we see a generator that enters the market as a result of clearing in the capacity auction. Their initial cash holdings are around -\$32.5 million because of the capital costs incurred. However, the net gain from capacity revenues and operating in the energy market begins to slowly pull the generator out of debt.



(a) Capacity clearing prices for all model years run.

(b) Combined cycle generator holdings from entry to end of model run.

Figure 3.13: Impact of capacity prices on generator holdings

The capacity payment sets the generator on track to make back its capital costs by the end of the 20-year payback period. After 5 years in the market, this generator has about $-24 million in debt. That's a rate of about $8.5 million every 5 years, which means that at the end of its 20-year life, it will have earned enough money to pay off its initial costs.

A total of 10 generators enter the system when the capacity market is in place. The increasing amount of combined cycle suppresses the more expensive combustion turbines, lowering the overall average price. Of course, once the approximately $13.5/MWh capacity price is added on, the advantage of having lower locational marginal prices in this case compared to the others disappears. The capacity payments made in this scenario amount to \$5.027B, more than half the amount paid to generators in the day-ahead markets in the same time frame.

Contrasted with the energy-only case, the expensive combustion turbines are dispatched much less. There is no loss of load, so there is no direct comparison to the energy only markets. A 0 day in 10 year loss of load probability is extremely good.

3.12 Comparison

Below are the results of the 5 scenarios run. All dollar amounts are in billions. We compare two methods of accounting, an optimistic version and a pessimistic version. The optimistic case serves as much load as possible, and only that extra demand that could not be met in each hour is shed. In the pessimistic case, if demand could not be met in a given hour, the entire load for that hour is dumped, equivalent to a blackout.

As can be seen from Table 3.9, the higher the allowable price for electricity, the fewer the outages incurred by the system. Looking at the price cap vs. hours lost, extrapolating, a price cap higher than \$20,000 would result in a single hour of lost load in 10 years. There is a local minimum in the optimistic case at a price cap of \$22,000 where the price cap is high enough to encourage new generation resources to be built but not so high that paying this price cap for an hour in 10 years results in extremely high payments.

Regardless of the price, this still requires load to be shed for these generators to be willing to enter. If the costs of the lost load can be absorbed by the system, then this would be preferable to the astronomical costs of the capacity market. Moreover, this model has deterministic demand. There is no guarantee that the one hour the new entrant needs will

	Base	$5,000	$10,000	$20,000	Cap Market
Hours Lost	168	78	29	2	0
MWh Lost	17,658	1,245	219	4	0
Optimistic					
Cost to serve allowable load	9.940	12.130	11.539	10.071	9.821
Total Value of Lost Load (@$9000/MWh)	0.159	0.011	0.001	0.000	-
Capacity Payments	-	-	-	-	5.027
Total	10.099	12.141	11.541	10.071	14.848
Pessimistic					
Cost to serve complete hours	9.895	9.843	9.835	9.836	9. 821
Cost of complete lost hours	8.853	4.128	1.535	0.106	-
Capacity Payments	-	-	-	-	5.027
Total	18.748	13.970	11.370	9.942	14.848

Table 3.9: Comparison of scenarios run

ever occur in the time period needed. There is a tradeoff between lower prices and more steady outages and a low probability event and a high spiking price. The capacity market is designed to mitigate any and all of that risk, allowing for a guaranteed payment.

A common measure of risk or reliability used by system operators worldwide is the concept of Loss of Load Expectation (LOLE). The most commonly implemented target LOLE is 1 day in 10 years, meaning that the ISO hopes to only have a single interruption in power delivery once in a 10-year period, on average. The earliest reference to the LOLE target of 1 day in 10 years was can be found in 1947, though the exact origin and determination of the target is unknown as mentioned in Milligan et al. (2011). Nonetheless, 1 day in 10 years has continued to be used as a benchmark for an acceptable level of risk.

The base scenario, in which the price cap is set just higher than the highest possible marginal cost at $47/MWh, results in 168 days of lost load over 10 years. Increasing the price cap decreases the amount of outages as expected; higher prices incentivize more generators to enter and ensures higher supply of generation in the market. A graph of the number of days lost relative to the price can be seen in Figure 3.14. We measure the risk as ?Days Lost.? The higher the level of risk we assume the more days of outages we face. There is a tradeoff here with the price cap. As opposed to an expected return, the price cap is a cost, so we invert the y-axis. A lower price cap represents a lower cost, or higher ?savings,? with the tradeoff of more risk. The higher price caps result in lower ?savings? but lower risk as well.

The $5,000/MWh and $10,000/MWh price cap case result in 78 and 29 days lost over 10 years, respectively. While these represent a significant improvement over the base case, they are a far cry from the widely accepted 1 day in 10 years. For the $20,000/MWh

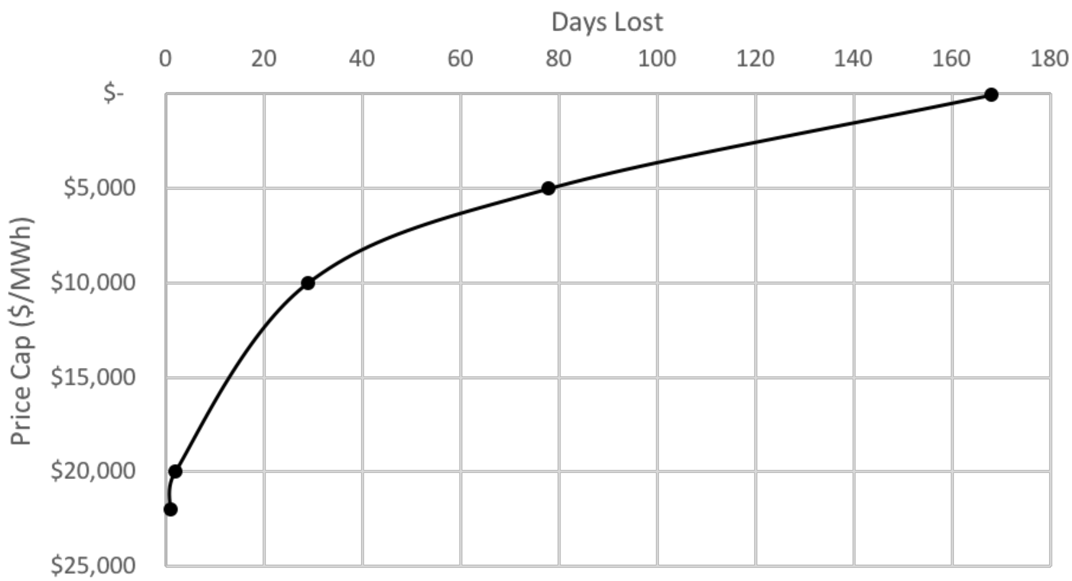
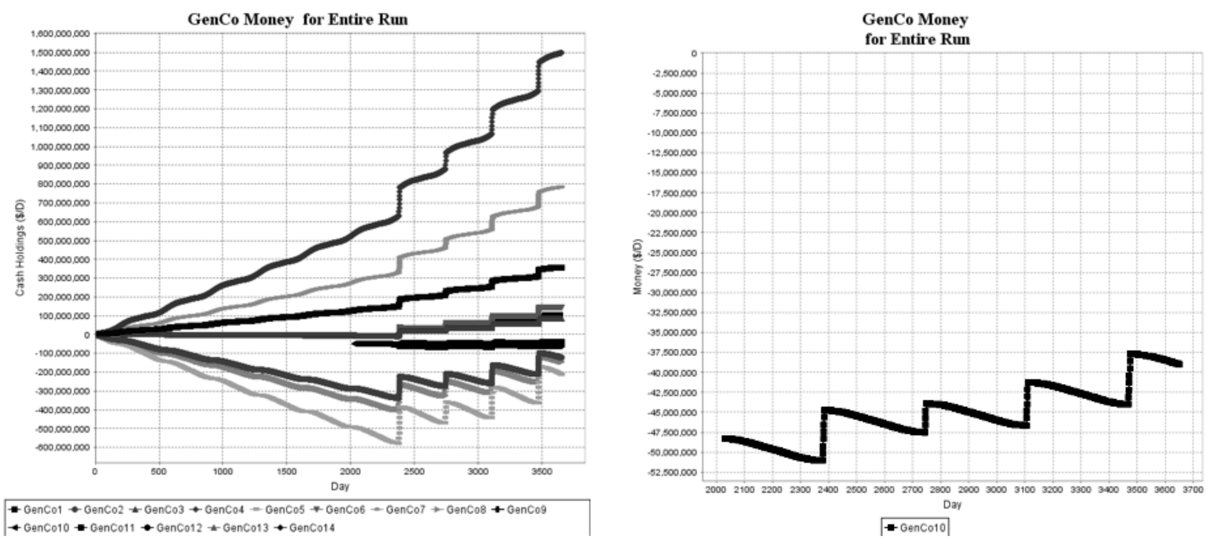


Figure 3.14: Days of Outages vs Price Cap

price cap scenario, the model achieves a 2 day in 10-year LOLE, a 15-fold reduction from the $10,000 case for a two-fold increase in the cap. The price cap that was found to result in the lowest total cost to the system, $22,000/MWh, is also the smallest price cap to achieve the 1 day in 10-year LOLE target, showing that it is possible to have lower costs of serving electricity as well as higher reliability (see Table 3.9, row 3).

It is hard to recommend the base case, as even though it is cheapest in the generous accounting method, having one's power go out every single day for a month straight every year would no doubt be problematic. The best option then is to remove artificial price caps put in place by system operators. As we have seen with ERCOT, the price cap has risen from $2,500/MWh in 2011 to $9,000/MWh in 2015, as ISOs realize that artificially suppressing prices reduces the incentives for new investment. A healthy market would have an uncapped price or at the very least a high price ceiling. The high prices would allow generators to recover their costs.

It should be noted here, however, that no new generators actually recovered their costs. Only in the capacity market did any generation resource (outside of coal) recover from their fixed operating costs. Recall that the size of the generation resources entering played a part in this. If instead we set mean capacity size for a new generator to 60MW, less generators enter: 4 compared with 10 at a price cap of \$20,000/MWh. Additionally, more hours are lost, 24 as opposed 2, and as a result more MWs go unserved.



(a) All generators (b) New combined cycle generator

Figure 3.15: Cash holdings when new entrants are 60MW plants.

Importantly for the generators themselves, capital costs are on track to be recovered by the 20-year period, as evidenced in Figure 3.15a. The large combined cycle generators in the bottom portion of the graph exhibit a sawtooth pattern that drifts upward. The new

		60 MW		30 MW	
Year	Peak Demand	Margin	Capacity	Margin	Capacity
1	5440	5.7%	5750	5.7%	5750
2	5500	4.6%	5750	4.6%	5750
3	5559	3.4%	5750	3.4%	5750
4	5618	2.3%	5750	2.3%	5750
5	5677	1.3%	5750	1.3%	5750
6	5737	0.2%	5750	0.2%	5750
7	5796	0.0%	5795	0.5%	5827
8	5855	0.0%	5853	0.4%	5880
9	5914	-0.2%	5902	0.2%	5927
10	5973	-0.3%	5955	0.4%	5998

Table 3.10: Reserve margin comparison with new entrants of size 60MW and 30MW.

combined cycle generator shown in Figure 3.15b enters with a -\$47.5M debt, but in four and a half years has earned \$10M, bringing debt to -\$37.5M. If the model were extended another 15 years, we would see the plant end the simulation with positive cash holdings. In the table of reserve margins compared below (Table 6), we see that the 60 MW generator case consistently has periods of outages, as the reserve margins are below 0% on the peak days of the year.

Furthermore, the 60 MW has a total cost of serving demand \$12.7B, \$2.5B more than the 30 MW generator size case with the same price cap of \$20,000/MWh due to the twenty-two additional hours of maximum prices.

Analyzing the impact on all stakeholders, we conclude that while this is good for generator revenues and operations, it is bad for the consumers of electricity. Having more outages and higher prices is not preferable to the scenario where generators entered and cannibalized their own profits. Of course, that situation is obviously not preferable for the generators, and, would likely not occur in real markets as generators would shut down after continually suffering losses year after year. Likewise, it is also unlikely that a single new generator would be able to serve all of the demand growth in a market. Regardless, let use the 60 MW case in our comparison with the capacity market scenario from earlier. The total costs for the energy only market were $12.7B and for the capacity market $14.8B. This averages out to an LMP of $40/MWh and $47/MWh. While a seven dollar difference seems high, consider that at the consumer level, where electricity prices are seen in dollars or cents per kilowatt-hour, this is a difference of 0.7 cents per kilowatt hour. For an average household using 11,700 kwh per year, this is an additional $82 per year for 0 hours of outages in 10 years. For generators, it is a stable and reliable source of income that guarantees recovery of capital and fixed operating costs.

Chapter 4

Learning

In this chapter we combine the theoretical work done in chapter two on optimal bidding strategies with the simulation model of chapter three. We have seen how the market behaves and performs when generators are bidding optimally (as defined in Chapter 2, where existing generation bids 0 and new generation bids their true cost of entry), we now wish to allow generators that are bidding into a capacity market the ability to use their own strategies for bidding. The question we address in this chapter is: If generators are not imbued with the optimal bidding strategy, will they learn how to bid optimally so that their behavior is consistent with Chapter 2 and the system outcomes are similar to Chapter 3?

A model in which participants learn about the game and the other players over repeated experiences is thus fitting. In such a game, a single decision maker acting in isolation is not applicable. There are other players in this strategic environment which affects the information gained by the individual. Furthermore, the information and experience gained by the other players lead to changes in those players' behaviors as well.

The relevant literature uses the term "learner" to describe an algorithm utilized by simulated decision makers that can be updated as the outcome of decisions is experienced over time. A suitable learner as described by Roth and Erev (1995) then has three main criteria. The first is that choices that lead to good outcomes are more likely to be chosen again in the future; this feature is referred to as reinforcement. Second, only the choices of other participants are available to an individual, not their entire strategy. Finally, the model should not depend on observations about the players that cannot be observed from the data itself. Though this may seem similar to the previous point, the distinction here is that while the second point dealt with strategies, the third point deals with attributes such as prior beliefs, or the updating of beliefs, or distribution of choices, for example. These are not modeled by the learner. Such a simple model allows for application to a wide variety of games without the need to fit parameters for each game. In this model, we use the same type of reinforcement learner for both capacity market auctions and day-ahead market auctions.

Here we focus on the application of the learner for capacity auctions. It meets all of the criteria outlined in the paragraph above. First, a good outcome in the case of a capacity auction would be receipt of a capacity payment that meets or exceeds the necessary payment required for new entry. Second, the choices of the other participants are observed by an individual in the forms of auction clearing capacity amounts and clearing prices. This is closely tied to the third point, in which we do not need to make any observations on the participants in the auction that would be considered private information, such as finances or true costs.

4.1 Roth-Erev Learning Algorithm

The basic model as described in Roth and Erev has each player i with a fixed number of actions, M_i. in the action domain M_i. At time $t = 1$, each player has an initial propensity to play the m^{th} action $q_{im}(1)$. A propensity is a real number, which, relative to the propensities on the other actions, determines how likely a player is to play that particular action.

So, if a player i plays action m and receives a payoff x (in dollars), the propensity to play action m is updated according to the formula

$$q_{im}(t+1) = q_{im}(t) + x. \tag{4.1}$$

Propensities increase as positive rewards are received. The probability that a player i plays action m is given by $p_{im}(t) = q_{im}(t)/\sum q_{ij}(t)$. The sum in the denominator is over all the actions $j \in M_i$. Over time, the rewards x have less marginal impact on the propensities and probabilities. Roth and Erev mitigate this by introducing a "forgetting" parameter into the model. This parameter $\phi \in [0, 1]$ is used to limit how flat the learning curve can become. By "forgetting" some prior experience, rewards will continue to have a significant impact on propensities over time. Each propensity is multiplied by $1 - \phi$ at the end of a period t, giving the new formula

$$q_{im}(t+1) = (1-\phi)q_{im}(t) + x. \tag{4.2}$$

The forgetting parameter also prevents the denominator $\sum q_{ij}(t)$ term from approaching infinity.

A second modification that Roth and Erev introduce is that of an experimentation parameter. This parameter prevents certain actions from dominating the domain space. It is possible that for low initial propensities, an extremely high payoff could skew the probabilities greatly, creating an instance where the player may never choose a different action from the one originally chosen. This would prevent the player from fully having a chance to learn about the decision context. To counter this, the experimentation parameter $\epsilon \in [0, 1]$ is introduced. For an action m the propensity is updated according to the formula

$$q_{im}(t+1) = q_{im}(t) + (1-\epsilon)x. \tag{4.3}$$

The remaining quantity ϵx is distributed evenly across the remaining actions' propensities

$$q_{ik}(t+1) = q_{ik}(t) + \frac{\epsilon}{M-1}x \quad \text{for } k \in M_i,\ k \neq m. \tag{4.4}$$

4.2 Generator Offer Curve Learning

Generators in the model have the option to engage in both economic capacity withholding and physical capacity withholding. Economic capacity withholding is the overstatement of marginal costs, while physical capacity withholding is the understatement of maximum generating capacity. Sun and Tesfatsion's AMES model features a modified version of Roth and Erev's reinforcement learning algorithm (Roth and Erev (1995); Erev and Roth (1998)). Each generator has a set of possible actions called an action domain. For simplicity and tractability, the action domain is finite. Each generator agent i has a propensity to choose a particular action m given by $q_{im}(t)$. The initial propensity is set exogenously to $q_{im}(1)$, and furthermore, $q_{im}(1) = q_i(1)$, that is, all initial propensity levels for actions are set equal for the first day of the model. Using the formulation of Roth and Erev, for a generator i who choses action m, the propensity is updated according to

$$q_{im}(t+1) = (1-\phi)q_{im}(t) + (1-\phi)x \tag{4.5}$$

and

$$q_{ik}(t+1) = (1-\phi)q_{ik}(t) + \frac{\epsilon}{M-1}x \quad \text{for } k \in M_i,\ k \neq m \tag{4.6}$$

In electricity markets payoffs can be negative as operating the power plant would lose money if electricity prices are lower than costs. Similarly, in capacity markets, a low bid might result a capacity payment that does not cover the net cost of new entry. While the propensity updating equations are unaffected by negative payoffs, the probability calculation for player i's action m,

$$p_{im}(t) = \frac{q_{im}}{(\sum q_{ij}(t))}, \tag{4.7}$$

are no longer valid probabilities if any propensity q_{im} is negative. While there are several possible solutions, we choose to modify the choice probability calculation to leave the reward function intact and easily parsed. In our model, negative rewards represent losses, which should not be shifted to become positive. In the AMES model, the probability that an action m is selected by generator i is given by

$$p_{im}(t) = \frac{\exp\left(q_{im}(t)/T_i\right)}{\sum_{j=1}^{M_i} \exp\left(q_{ij}(t)/T_i\right)} \tag{4.8}$$

The parameter T_i is a cooling parameter that affects the impact propensities have on Generator i's choice probabilities. If T_i is large, then the choice probabilities go to $1/M_i$, and if T_i is small then the choice probabilities are heavily loaded over the choices with the highest propensities in the domain. Exponentiation ensures that any negative propensities are transformed to positive values, giving positive choice probabilities. The model with the cooling temperature and exponential probability calculation is a variant of the original Roth-Erev reinforcement learner and is referred to by Sun and Tesfatsion as the VRE-RL algorithm. All references to learner henceforth refer to the VRE-RL algorithm used in the AMES model.

4.3 Capacity Offer Learning

Using the modified Roth-Erev learning model described above a learning component will be added for the capacity bidding in the yearly capacity auction. The action domain will be a percentage α_i in the range $\alpha_i \in [-1, 1]$, with a finite number of increments. This parameter α_i is the factor by which each generator adjusts and reports their levelized cost of electricity C resulting in a bid of $(1 + \alpha_i)C$. Recall from Chapter 3 that the levelized cost of electricity (LCOE) was defined as

$$LCOE = \frac{CapCost + FOM}{ExpectedGeneration} + VOM + Fuel \tag{4.9}$$

For new entrants, there are multiple ways to implement capacity bid learning, each with positives and negatives.

We devise an approach where each new entrant belongs to one of a set of holding companies. The approach here is meant to represent a company having a portfolio of generation, both existing and prospective. Each holding company therefore has a learner for their existing generation, and one for new plants they are planning to submit into the capacity auction. The number of new entrants in a given year is fixed and the number of prospective new entrants is less than the number of holding companies. Each new entrant is assigned to a holding company and has access to the new entrant learner of the holding company.

Every year, every action chosen by prospective new entrants of a holding company are drawn from the same distribution. This does not mean that they choose the same action, just that they have access to the same body of knowledge and propensities. After the conclusion of the auction, each of these generators would provide information back to the single learner of their holding company. We are able to gather information at a faster rate than if each entrant had its own learner.[1]

Should any generators clear in the auction, they are moved into the pool of existing generators, and after updating the holding company's new entrant learner, are removed from that pool. Instead, they now choose actions and update propensities based on that same holding company's existing generator learner. Those that do not clear must wait until next year's auction to try again. The number of prospective new entrants is fixed every year, so new generators enter the pool to replace those that cleared in the auction. The manner in which each learner processes the results of the auction is discussed next.

4.3.1 Signal Updating

At the conclusion of a capacity auction, the clearing price is revealed. Each generator then updates the propensity q_i for action α_i with the resulting reward. Additionally, the clearing price provides more information than just whether or not the generator's particular bid cleared. The structure of the market also allows it to provide information on what reserve margin cleared. This will allow us to update more than one action α_i per auction, as we will now describe in more detail.

Upon model initialization, N new entrants are created with normally distributed random generator capacity size and annualized cost C. All are assigned to one of H holding

[1]Other learners were tested. One of which was a single global learner. In this scenario, for any given year there are a fixed number of entrants. Every entrant shares the same learner. Each auction allows for fast learning on account of there being a single learner with multiple updates per auction. In tests, the shared learner resulted in quick convergence, however the idea that every prospective generator shares the same knowledge is not very credible. A second option was that of each generator having its own learner. This results in slow convergence in the time frame allowed by the model and computational resources.

companies h_i. All entrants have the same initial propensities over the actions in the action domain, and thus equal probabilities. When the auction begins, each new entrant submits a bid equal to $(1 + \alpha_i)C$. All bids are sorted and stacked to form a supply curve, then intersected with the demand curve set by the ISO. The clearing price is paid to all entrants whose bids were to the left of the intersection point, in exchange for a commitment to be available for unit commitment in the day-ahead and real-time markets. Each entrant's propensity is updated with the appropriate reward (0 if not clearing the auction and clearing price minus fixed cost $p^* - C$ if clearing the auction) for the action α_i they chose. If k of the new entrants cleared, then they are removed from the new entrant pool, and k new entrants are added to the pool, again with randomly distributed capacity and size, and the same initial propensities and randomly assigned holding company. The entrants moved out of the new entrant pool are marked as existing generators, and use the holding company's existing generator learner as opposed to the holding company's new entrant learner. This process continues until the simulation is over. It is possible that a new entrant never clears in any of the capacity auctions and is never replaced.

4.3.2 Low Bids

To increase the amount of learning occurring in a given year's auction, additional signals were considered. If a bidder does not clear in the auction they cannot observe other bidder's bids, but they can observe the clearing price. Let us introduce a new price p' that is equal to the price given by the demand curve for the quantity of cleared capacity plus the additional capacity of Generator i.

$$p' = D(cap_{cleared} + cap_i) \tag{4.10}$$

Each losing bidder updates each action m that would have resulted in a bid lower than the clearing price ($(1+\alpha_i)C < p'$) with the reward for clearing in the auction, $p' - C$. In other words, this assumes that, had the bidder bid something lower than their original bid, they would have cleared in the auction and received the clearing price p' and paid cost C. More often than not, this results in negative rewards on extremely low bid factors. This result is not unreasonable; at the extreme, one would assume that bidding a cost of 0 when one's true cost is nonzero would not be in one's best interests.

4.3.3 High Bids

Similarly, consider the case of every bidder whose true cost was less than the clearing price $C < p^*$. Every action i that would have resulted in a bid higher than the clearing price $(1+\alpha_i)C > p^*$ is given the reward $-(p^* - C)$. Since $p^* - C$ is positive, this negative reward is meant to represent the opportunity cost of bidding too high, not clearing in the auction, and having to wait another year for the next auction. By failing to clear in the capacity market, the generator is missing out on potential revenue that could have been earned that year.

4.3.4 Examples

The examples presented below will be broken into two parts for clarity. The first part will provide a simple example of the Roth-Erev learning algorithm in action, and the second part will provide an example of the reward functions described in the capacity offer learning section.

Consider this basic application of the Roth-Erev learner with a simple ultimatum game for two players. There are 10 dollars available to split between two players. Only integer denominations are allowed. One player is assigned the role of offer maker (player one), and one player the role of offer taker (player two). Player one presents a dollar amount for themselves to keep to player two, and if player two accepts, player two gets the remaining funds. If player two rejects the offer, then both parties get 0. Player one begins with a demand d_1, an integer between 1 and 9 inclusive. Player two reports a maximum allowable demand m_2 that is also an integer between 1 and 9 inclusive. If $d_1 \leq m_2$ then player one receives d_1 and player two receives $10 - d_1$. Otherwise, if $d_1 > m_2$ then both players receive 0.

We will set up an example where player one always has the role of offer maker, and player two that of offer taker. The game will be repeated multiple times to show the effects on propensities and action probabilities. We will leave experimentation at 0 and recency at 0, and all initial propensities on actions are 1 as well. For each player there is an action domain of size 9 (integers 1 through 9) which we will label a_{ki} where i is the integer chosen and k is the player number. Assume in the first game that player one chooses $d_1 = 4$(action a_{14}) and player two chooses $m_2 = 6$ (action a_{28}). Then player one receives 4 and player two receives $10 - 4 = 6$. The propensities are then updated for both player one and two on the respective actions that they chose. In player one's case

$$q_{14}(1) = q_{14}(0) + 4 = 1 + 4 = 5, \tag{4.11}$$

and in player two's case

$$q_{26}(1) = q_{26}(0) + 6 = 1 + 6 = 7. \tag{4.12}$$

Now player one has a higher chance of playing action 4 in the next period, and player two has a higher chance of playing action 8 in the next period. The probability that player two plays action 8 is higher than the probability that player one plays action 4, since the propensity is higher. The other actions propensities remain at their initial values of one.

Now we reintroduce the recency and experimentation parameters ϕ and ϵ. We will let recency be $\phi = 0.5$ and experimentation $\epsilon = 0.25$. If we take the exact same game and moves as before, the difference in propensities is now

$$q_{14}(1) = (1-\phi)q_{14}(0) + (1-\epsilon)4 = (0.5)(1) + (0.75)(4) = 3.5 \tag{4.13}$$

for action 4, and

$$q_{1j}(1) = (1-\phi)q_{1j}(0) + \frac{\epsilon}{8}4 = (0.5)(1) + \frac{(0.25)(4)}{8} = 0.625 \tag{4.14}$$

for all other actions $j \neq 4$. Similarly, for player two:

$$q_{26}(1) = (1-\phi)q_{26}(0) + (1-\epsilon)4 = (0.5)(1) + (0.75)(6) = 5 \tag{4.15}$$

for action 6, and

$$q_{2j}(1) = (1-\phi)q_{2j}(0) + \frac{\epsilon}{8}4 = (0.5)(1) + \frac{(0.25)(6)}{8} = 0.6875 \tag{4.16}$$

for other actions $j \neq 6$. In a situation where the maximum acceptable demand is lower than the amount demanded, both players get nothing. The reward payout is 0, and in models where recency ϕ is zero, there is no change to the propensities and thus no change to the probabilities. When recency is non-zero, the propensities are reduced by a factor of ϕ, which results in a change in the probabilities.

Let us now consider the extra signal information described in the low bids and high bids sections above. We consider the same two player offering game with a few slight modifications and are only concerned with the reward function for player two. Now, rather than setting a demand for their own payoff, player one offers an amount o_1 for player two to receive. Player two has a minimum acceptable offer f_2, and if this minimum is less than the amount offered by player one, player two receives the offered amount. Furthermore, any amount received by player two is subject to a "gambler's tax" of 4 dollars. The reward function then for player two is $o_1 - 4$ if $f_2 \leq o_1$ and 0 otherwise. Suppose that player one plays action 6 ($o_1 = 6$) and player two plays action 3 ($f_2 = 3$). The updated propensity for player two's action 3 is

$$q_{23}(1) = q_{23}(0) + 6 - 4 = 1 + 2 = 3 \tag{4.17}$$

To utilize the extra information, player two now considers the other bids that could have been made: the low and high bids discussed previously. For the "low bids", all minimum acceptable offers that would have been less than $o_1$are updated. These other offers are actions 1,2,4,5 and 6.

$$q_{21}(1) = q_{21}(0) + 6 - 4 = 1 + 2 = 3 \tag{4.18}$$

$$q_{22}(1) = q_{22}(0) + 6 - 4 = 1 + 2 = 3 \tag{4.19}$$

$$q_{24}(1) = q_{24}(0) + 6 - 4 = 1 + 2 = 3 \tag{4.20}$$

$$q_{25}(1) = q_{25}(0) + 6 - 4 = 1 + 2 = 3 \tag{4.21}$$

$$q_{26}(1) = q_{26}(0) + 6 - 4 = 1 + 2 = 3 \tag{4.22}$$

For "high bids", all minimum acceptable offers that are higher than o_1 are considered. In this case, that would be actions 7,8, and 9. The goal behind updating these actions is to consider whether or not playing a lower minimum would have resulted in positive reward. If that would be the case, then the propensities are updated with the opportunity cost of having missed that reward. An example should make things clearer.

$$q_{27}(1) = q_{27}(0) + -(6-4) = 1-2 = -1 \tag{4.23}$$

Here, the reward received had a lower action been played would be 2 (6-4), but because player two was "greedy" they missed out on receiving this reward, hence the -2 in the equation above. The same logic applies to actions 8 and 9.

$$q_{28}(1) = q_{28}(0) + -(6-4) = 1-2 = -1 \tag{4.24}$$

$$q_{29}(1) = q_{29}(0) + -(6-4) = 1-2 = -1 \tag{4.25}$$

We conclude this section with one final example that illustrates one interesting (and common case in our capacity markets). We will continue the previous example, so we are now in period 2. Here the offer played is $o_1 = 3$ and the minimum acceptable offer chosen is $f_2 = 4$.

$$q_{24}(2) = q_{24}(1) + 0 = 3 + 0 = 3 \tag{4.26}$$

Because the minimum acceptable offer was greater than the offered amount, the reward received was 0 and the previous propensity from the last period is carried over with no change.

For a lower minimum acceptable offer (actions 1, 2, and 3) the amount offered is received and subject to the "gambler's tax", thus resulting in net gain to player two of -1

$(= 3 - 4)$. The propensities would be updated thusly:

$$q_{21}(2) = q_{21}(1) + (-1) = 3 - 1 = 2 \tag{4.27}$$

$$q_{22}(2) = q_{22}(1) + (-1) = 3 - 1 = 2 \tag{4.28}$$

$$q_{23}(2) = q_{23}(1) + (-1) = 3 - 1 = 2 \tag{4.29}$$

Higher minimum offers would not have resulted in a positive reward (the reward will always be $3 - 4 = -1$), and so those propensities are not updated according to our high bids rules.

$$q_{25}(2) = q_{25}(1) = 3 \tag{4.30}$$

$$q_{26}(2) = q_{26}(1) = 3 \tag{4.31}$$

$$q_{27}(2) = q_{27}(1) = -1 \tag{4.32}$$

$$q_{28}(2) = q_{28}(1) = -1 \tag{4.33}$$

$$q_{29}(2) = q_{29}(1) = -1 \tag{4.34}$$

At the end of this step we have actions 4, 5, and 6 as the highest propensity actions, and 7, 8, and 9 as the lowest. We note that the actions 1, 2, and 3 are lower due to the gambler's tax, and those actions should not be played as part of an optimal strategy. Were this exercise to continue, we would see the gap in propensities between the more mid-range actions and low minimums widen.

4.4 Learning results

The model setup for the results that follow is now described. There are two holding companies and a total pool of 10 prospective generators per year. This allows for a balance between faster convergence and the ability to run multiple learners in parallel. There is a 10 year "burn-in" period in which the simulation runs unrestricted, after which all results with the exception of the holding companies' propensities are wiped out. These propensities can be thought of as priors. The simulation then begins anew, as if on day 0, with initial demand levels and only the initial existing generation described in Chapter 3 and runs for 10 years. The reasoning behind the burn-in period is that in preliminary testing, more than 35 generators would be added in the first 5 years of a twenty-year simulation. Generators were bidding wildly, having started with equal probability of every action. With such a massive buildup of excess capacity in the early periods, the consequent learning results and their applicability to energy markets in the United States must be questioned. The simulation no longer represents a realistic energy market as there would not be situations where a market develops a 100% reserve margin. The decision was made to take 10 years of this learning and use it as prior propensities for the learners in a 10-year simulation immediately following, to more align with and better compare to the models run in Chapter 3.

The forgetting parameter is set to 0.2 so that the reward received from an action is added to 80% of the existing propensity for that action. This means that rewards associated with the oldest actions chosen are worth less and less as time goes on. In early years, since the actions start with equal probability of being chosen, there is a high amount of variance in the rewards. The agents are still learning the best action or actions to play. As a dominant strategy emerges in later years, however, the weight given to these early actions and rewards should be lessened. The early actions can be thought of as feelers or test bids before more educated bids can be made. We slowly phase out their contribution in favor of rewards from more informed actions. Experimentation is set to 25%, such that 75% of the reward gets applied to the chosen action and 1.25% (25%/20) gets applied to every other action. Experimentation is not set too high so as to allow for a dominant strategy to emerge, but not so low that a local minimum cannot be exited.

4.4.1 Optimal Bid in Learning Case

There are two bid profiles to consider, both with different optimal strategies. One is the bid profile of new entrants, and the other the bid profile of existing entrants. Recall from Chapter 2 that it is in the best interests of existing generators to bid 0, as anything higher results in a potential loss of capacity market payments. Because the generator is already online; costs are sunk. Should the generator not clear in the auction, they do not have the option to "give back" the generator, or to not pay that years' capital cost. It is in their best interest to try and receive any capacity payment, regardless of the amount, which would mean bidding as low as possible. Attempting to clear at a higher price by submitting an inflated bid runs the risk of not clearing at all.

The new entrants optimal bid is their true cost. Any lower and they too do not recover the entirety of their capital costs. Any higher and they have a lower probability of clearing in the auction and receiving nothing at all, except a year of waiting and the opportunity costs associated with the waiting. In this case that represents the lost income from entering and operating in the energy market.

We first run a case where we only "turn on" one set of the learners: the new entrants. We set existing generators to always bid 0, and the prospective generators use the learner described in the previous section. This scenario is run for 10 years (after the 10 "priming" years). No generator ever enters in any of the capacity auctions that are held. However,

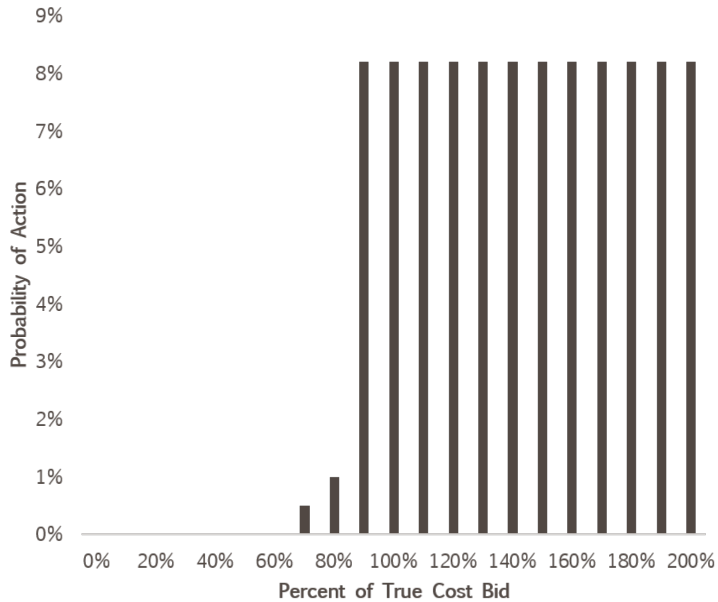


Figure 4.1: Averaged probabilities for different bids

nine generators are built though out the 10-year period as a result of favorable economic conditions. A big part of these favorable economic conditions comes as a result of predicted capacity payments. Recall that the LACE calculation takes into consideration both the predicted peak demand and the current level of installed capacity. With the two numbers, a reserve margin is calculated and from the reserve margin the capacity price that would arise from the demand curve set by the ISO.

Figure 4.1 is the result of the average of the holding companies for one such run. Because the new entrants will bid almost entirely at or above the true cost of entry, the probability of clearing in the auction with this strategy is very low.

As seen in Table 4.1, The reserve margin always hovers around 1.5%, with new entrants coming online as reserve margin drops lower and predicted capacity price rises. We

Year	Reserve Margin	Installed Capacity	Peak Demand
1	5.6%	5750	5446
2	4.5%	5750	5505
3	3.3%	5750	5564
4	2.2%	5750	5624
5	1.7%	5778	5683
6	1.7%	5838	5742
7	1.6%	5896	5801
8	1.5%	5946	5861
9	1.6%	6016	5920
10	1.7%	6081	5979

Table 4.1: Reserve margins for the case when only prospective generators learn

know the new entrants are coming online as a result of predicted capacity prices because in the scenarios with no capacity market from Chapter 3, prices do not get high enough in the electricity markets (when they are not allowed to spike to $20,000/MWh) to alone meet total costs. We see that the capacity prices stabilize at the price of around $12.5/MW-day after year six in Figure 4.2.

Figure 4.3 below shows a graph of the levelized cost of electricity and levelized avoided cost of electricity for the 10 years (non-benchmark) for a 30MW combined cycle generator. When LACE is greater than LCOE a new entrant enters the market. The gap between LACE and LCOE is what we use as our proxy for the net cost of new entry (CONE) that is the basis for all capacity payments. The sawtooth shape in the graph can be explained in a few steps. The initial gentle slope of the LACE is revenues from a market in which the only payments made to generators are the revenues earned in the day-ahead market. The first year has no capacity payments since a capacity auction has not been held. In the second year, because the initial model state has excess capacity, the predicted capacity price is 0

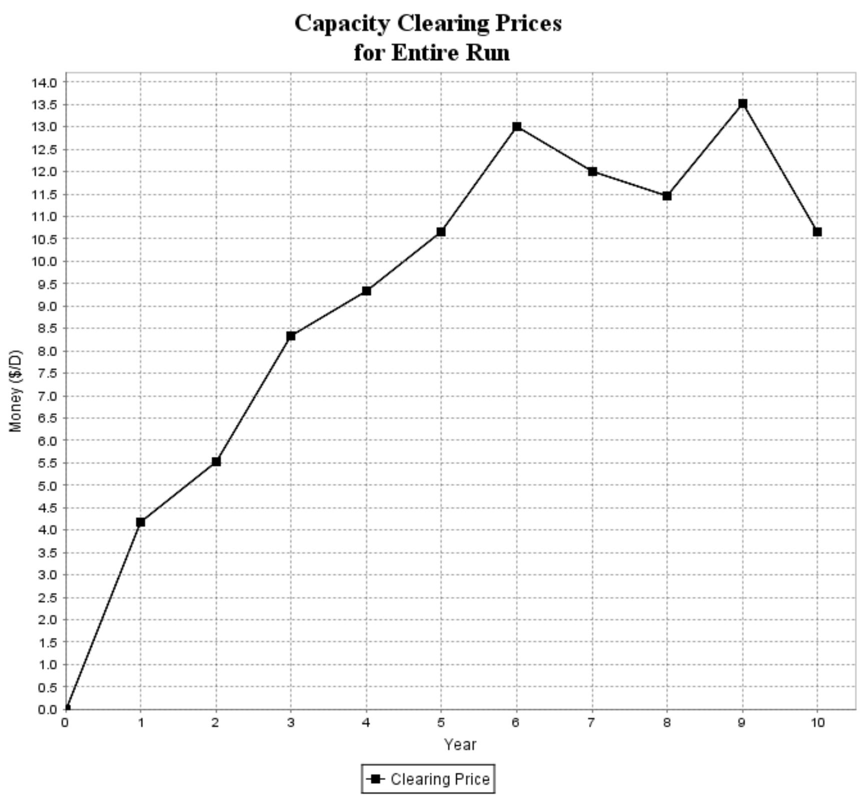


Figure 4.2: Capacity prices in the case when only prospective entrants learn

dollars, and the markets in general function as they did in the first year. The gentle rising slope is a result of demand growth and in turn higher prices due to heightened capacity utilization. We see the first portion of increased slope when the predicted capacity price is nonzero around day 730, or year 3. Around this time demand is high enough that the initial capacity existing in the model is not sufficient, pushing the reserve margin downward. It is also the first time predicted reserve margin drops below 5% and with it a non-zero capacity price is predicted.

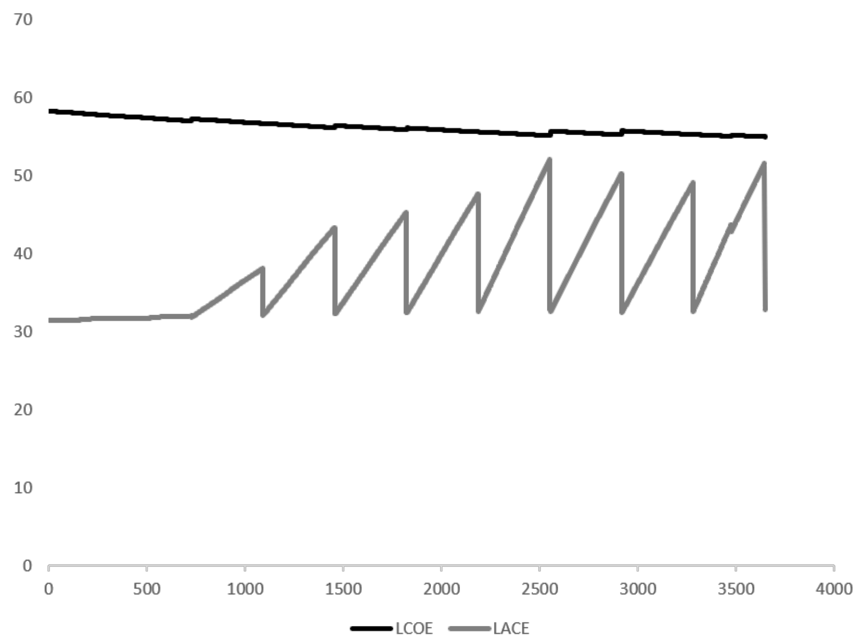


Figure 4.3: LACE and LCOE comparison in the case where only prospective generators learn

The drop that is seen around day 1099 is not because new capacity was built, but because a new year begins. A capacity auction is held on the first day of every year, and those that clear in it are paid the clearing price for that year. When calculating the LACE for the next year forward, a new entrant entering after the capacity auction, on day 2 for example, has 363 days of no capacity payments received, before they can bid into the auction the following year.

In subsequent years the LACE rises enough that the predicted capacity price, coupled with the days that price will be received, is sufficient to encourage new entrants 30-60 days before the start of a new year, on average. New capacity not only lowers predicted capacity prices but also lowers LMPs and thus generator revenues. Each year after the capacity auction the LACE value drops and slowly works its way back up again. There is a small kink in an otherwise smooth line around day 3478, where a generator enters the market through competitive entry rather than by clearing in the capacity auction. Also note that the two peaks with the highest LACE correspond exactly with the highest capacity prices, as expected.

All generators in the market are able to begin paying off their total costs, both variable and fixed, from the capacity payments received in the auctions. All existing generators clear in the auction with 0-dollar bids and the capacity price is always positive. Most, if not all (due to variability in initial capital costs), generators are on track to make back their initial construction costs before the end of their assumed 20-year lifetime.

4.4.2 Existing Capacity Bids According to Learned Strategy

If we now also let the existing generators learn how to bid in the capacity auction using the same format as before the results of the two learners look very different. A slight modification was made to the base set up of the model. The nine existing generators were each split into two equivalent generators, each half the size of the original and with the same cost structure. The reason for this is that when one of the very large generators failed to clear in the auction, this had the adverse of effect of at times removing more than 15% of capacity from the auction. The resulting reserve margin and clearing price led to scenarios where excessive amounts of new generation were built, causing a massive oversupply of generation resources. Though the doubling of the number of initial generators increased run time, the model more closely approximates both the theorized scenario in Chapter 2 and real-world scenarios. Observing the holding companies' propensities for existing generators at the end of the model run, and, averaging each action propensity across the entire set of runs, we see in the figure below that 99% of the mass is on bids in the range $\alpha \in [-1, -0.6]$, or a strategy of 60% or less of one's true cost.

The average probability on actions taken from 25 runs and 2 holding companies per run is given by the probability mass function in Figure 4.4.

For new entrants, a different picture is painted. Here, bidding low increases the chance of clearing in the auction, with the caveat that, if cleared, the generator is expected to be online and entered into the day-ahead market. Because the capacity payment is designed to meet the difference between the cost of new entry minus the payments from participating in the energy market, a low bid can result in the difference not being met. For example, say that it costs \$100,000/MW-year for a new combustion turbine. It can be expected to make \$25,000/MW-year from participating in the energy and ancillary services market. Thus, \$75,000/MW-year must be covered by capacity payments. If, instead the generator bids in at \$40,000/MW-year, and the market clears at that price, there is still \$35,000/MW-year that is missing. In Chapter 2, recall that the optimal bid for new entrants was just shy of truthful bidding.

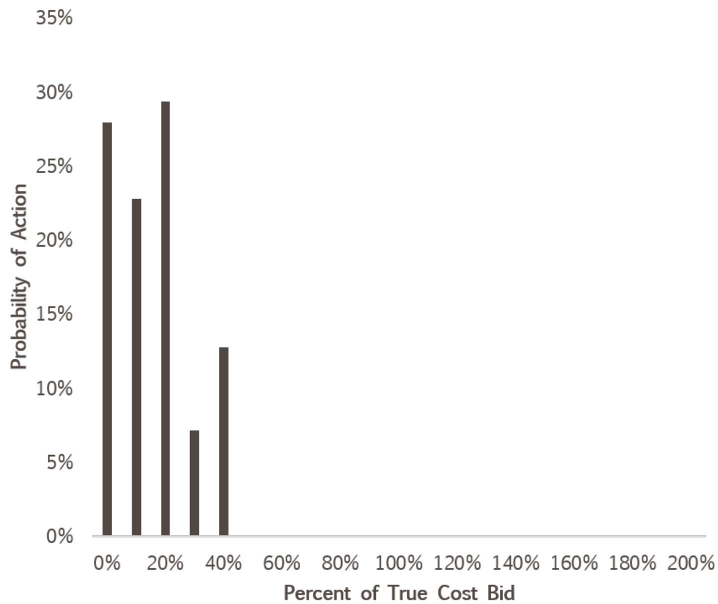


Figure 4.4: Averaged probabilities for existing generator learners

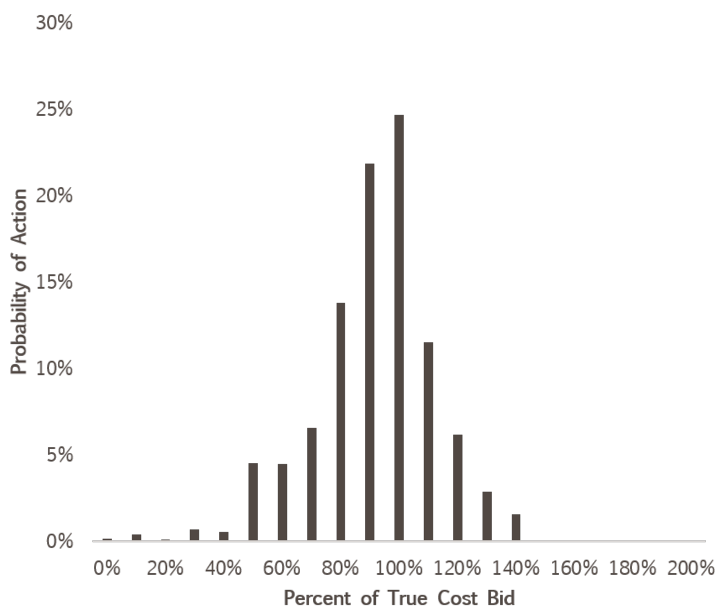


Figure 4.5: Averaged action probabilities for new entrants when all parties learn

Year	Reserve Margin	Installed Capacity	Peak Demand
1	5.6%	5750	5446
2	6.2%	5846	5505
3	5.5%	5872	5564
4	5.5%	5932	5624
5	4.4%	5932	5683
6	3.3%	5932	5742
7	2.3%	5932	5801
8	1.2%	5932	5861
9	1.4%	6001	5920
10	1.3%	6059	5979

Table 4.2: Reserve margin for market when all parties learn

From the distribution in Figure 4.5, we see that result is backed by simulation. It is not advantageous to bid low, as the capacity payments are not high enough to be rewarding. There is high mass on both truthful bidding and just shy of truthful bidding. The two combine for almost 50% of the entire mass function.

How does this affect the market? The usual metrics we look for, such as reserve margin and capacity clearing prices, look very similar to those in Chapter 3 as evidenced by Table 4.2 and Figure 4.6. The reserve margin hovers just above the target of 0%, at 1%. The reasoning for this being that the demand curve defined by the ISO is set to pay out the net CONE when reserve margin is 1% above the target level. Since the 0 bids from existing generators and the truthful bids from prospective generators create a situation where there is no supply being withheld, and no opportunistic entry, the equilibrium point stabilizes to where the capacity market covers the difference when necessary.

The capacity clearing prices for the first three years are high as the cleared reserve margin is less than 0%. In later years however, the price settles around \$13. An interesting

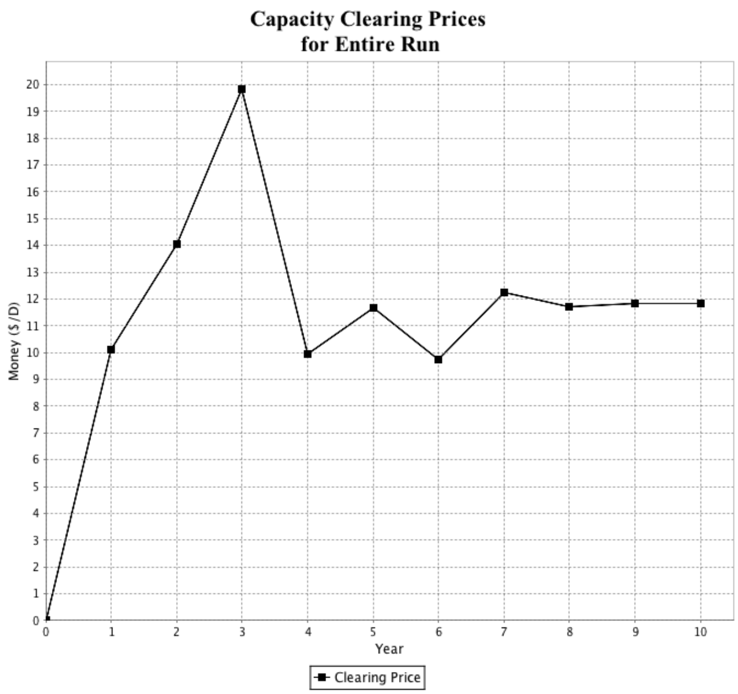


Figure 4.6: Capacity clearing prices when all parties learn

Year	Capacity Price	Cleared Capacity	Installed Capacity
1	10.12730	5626	5750
2	14.05059	5597	5846
3	19.84752	5407	5872
4	9.94348	5812	5932
5	11.67052	5832	5932
6	9.75118	5932	5932
7	12.26465	5932	5932
8	11.71958	6001	5932
9	11.84386	6059	6001
10	11.85249	6119	6059

Table 4.3: Mismatch when cleared capacity is lower than installed capacity

thing to note is that at the start of the model, the cost of new entry (which is based off of the difference of LACE and LCOE) for an average sized combined cycle generator is roughly \$13/MWh. Observe in Figure 4.6 that in the third year the clearing price is close to the maximum of 1.5 times the CONE. This means that the cleared reserve was significantly less than the target level. Even though there exists sufficient installed capacity in the market, not all of it clears in the auction; existing generators' bids in the first few years have not yet converged to the optimal strategy, resulting in bids that are too high. Table 4.3 highlights this disparity between installed capacity and cleared capacity.

We see evidence of the downside of not bidding 0 (shown to be optimal in Chapter 2) as an existing generator in Figure 4.7 below. Two generators have entered the market through projected LACE > LCOE.

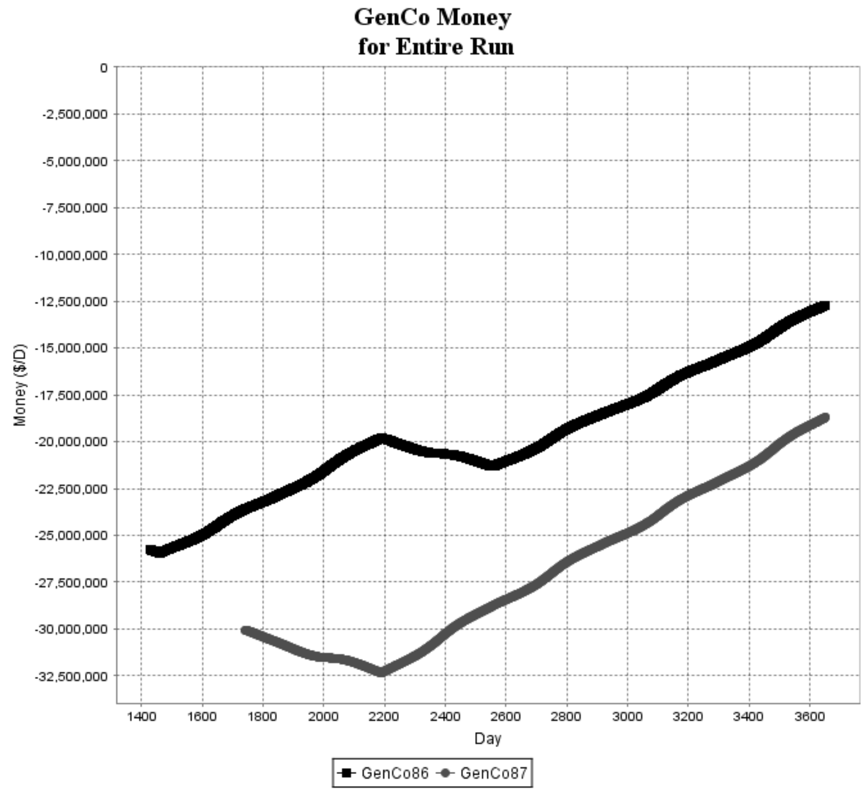


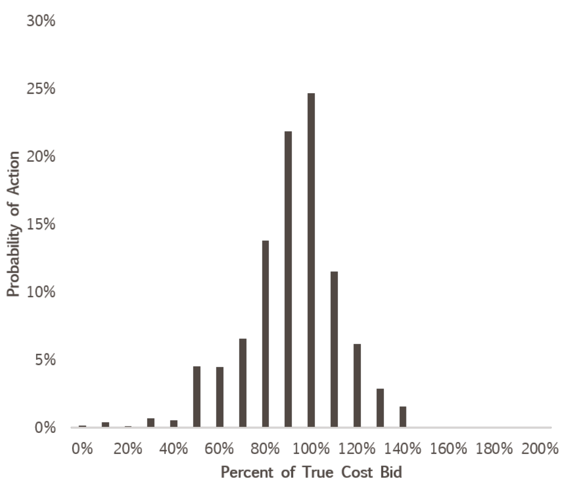
Figure 4.7: Generator cash holdings for two selected new entrants

The generator represented by the red line loses money at first but then, by clearing in the subsequent capacity auction, begins receiving capacity payments. The generator represented by the blue line enters but fails to clear in the auction held on day 1825. The cash holdings continue to decline until the second auction held after their entrance date, at which point they clear in all subsequent auctions and continue to earn a profit. We can all see the red generator fail to clear in that same auction, at which point they operate at a loss until the next year.

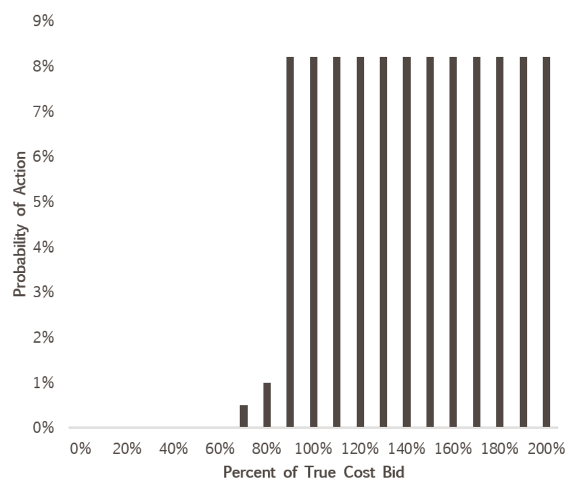
4.5 Learning Takeaways

When allowing generators to learn bidding strategies in a basic capacity auction, the results corroborate what was theorized in Chapter 2 and implemented in Chapter 3. For new generators, we see that it is optimal to bid truthfully. Doing so minimizes losses and opportunity costs of waiting. For existing generators, the lower the bid, the better. For existing generators, it is optimal to bid 0. Otherwise the fixed costs are never recouped under standard operating conditions. Any capacity payment, no matter how small, is better than none.

Comparing the final probability distributions of new entrants when existing generation is allowed to learn and when it is not we note that the actions are more closely centered around truthful bidding in the former case.



(a) Existing generation learns



(b) Existing generation does not learn

Figure 4.8: Comparison of action probabilities for new entrants

The reason for this is that when the existing generation cannot learn, the high bids rarely result in in offers that clear. The propensities mostly go untouched, as they did in simple example presented earlier in this chapter. Hence the rather uniform distribution across those actions above truthful bidding. When existing generators are allowed to learn, the capacity clearing in the auction no longer defaults at a minimum to the installed capacity in the market. Now the existing generation too has a chance to bid too high and not clear. Because of this, there is a higher chance of higher bids from new generation clearing. This also means that the opportunity cost penalties are more variable. Before, since reserve margins were always near the target, truthful bidding was the maximum allowable bid, and everything higher would incur the penalty. Now, a higher bid can still clear, essentially shifting the bids that receive that penalty.

From the perspective of the ISO, existing generation bidding in at 0 helps curb capacity payments, as without the base existing generation supply for "free", we see the wild spikes due to insufficient reserve margins when large plants do not clear. So too, we assume, would generators prefer the smoother, more predictable capacity payments seen in the simulation than the variability in prices exhibited in real life capacity auctions such as those of PJM.

Chapter 5

Discussion and Further Research Directions

The subject of capacity markets has been a hotly studied issue in electricity markets. Questions as to their viability or overall beneficence are debated by industry professionals and by academics alike. Though this analysis seeks only to clear a small portion of the haze surrounding capacity markets, it is our hope that the research can cast some light on the auction formats and provide a useful tool for evaluating different scenarios. In the theoretical framework of Chapter 2 we analyzed variations on the multi-unit, discriminatory supply auction. Starting from a single bidder playing against "the market," working our way towards the multiple bidder multiple winner demand curve-based auction. When modeled as a single player playing against a random capacity clearing price, it was optimal for the generator to bid in their true cost in the first period. For each subsequent period after the first period in which they clear in the auction, the optimal strategy is to bid 0. At this point, it does not matter what the clearing price is, the generator has to pay its capital costs and any amount of capacity price will help to offset those.

From there we consider single period games where multiple bidders compete for the right to supply one unit of generation into the market. Here we found that generators were likely to shade bids at every cost. Unsurprisingly, the more bidders present, the more severe the level of bid shading, as every generator participates in a race to the bottom to be able to supply the one unit of generation. The generator costs that result in 0 bids varies with the number of participants in the auction, but asymptotically approaches the mean of the price distribution. That is to say, when one's cost is below the mean of the unknown capacity price, bid 0, otherwise, bid one's cost adjusted downward relative to the distance from the mean price.

Acknowledging that having a single winner is not how the majority of capacity auctions work, we extend the multiple bidder model to multiple winners. A single price is paid to all winners, and the number of winners is known prior to the auction. As before, when holding the number of winners fixed, increasing numbers of participants lead to increased levels of bid shading. 0 bids are again apparent, with the same asymptote of 0 bids for any cost below the mean price. An interesting, but perhaps not unexpected result is that if the number of bidders is held fixed, but the number of allowable winners is varied, the higher the number of winners the less the degree of bid shading. Intuitively this makes perfect sense, there is less need to aggressively underreport your cost if you are likely to win just by participating.

Lastly, we modify the number of winners and clearing price to be dependent on the bids submitted. This is closest to the auction structure that is used by many ISOs in their capacity auctions. In this auction, in contrast with the others studied, the slope of the bid as a function of cost has the same slope regardless of the number of participants. What changes with an increasing number of participants is the translation of the biding function to the right, resulting in more 0 bids. Relative to our other bid functions, the optimal bid functions were much closer to truthful bidding.

In comparing the different auctions to each other we note a few key differences. When a flat price is paid to a fixed number of winners there are higher costs to the system, but more capacity procured to go with it. On the other hand, the decreasing price auction does not acquire as much generation capacity, but pays lower costs per unit overall. The lower levels of procurement could lead to outages, however we found that only at very high values of lost load ($100,000/MWh) did the cost of an outage make the flat price auction cheaper in terms of average expected costs.

We then took our theoretical analysis and applied it to an agent-based simulation model of a day ahead electricity market, where each generator is modeled by an agent. An approximation of the Texas energy market was constructed, and the simulation run over various parameters to test sensitivities. Additionally, the model was extended to include both market-based competitive entry and capacity auction-based entry. In all scenarios with an energy-only market, the system experienced multiple consecutive days of load, sometimes for multiple hours. By increasing the maximum price cap, these lost hours are able to be reduced. Perhaps counter-intuitively, it is not possible to reduce the lost hours to 0; for the price to hit the cap, there must be some constraint or outage. Generators deciding to enter the market factored this into their calculations, knowing that they would recoup a good portion of their costs for that single hour of high prices. Consider the cost of serving electricity and the cost of load together, we found an optimal price cap that resulted in the minimum possible total costs to the system.

With the capacity market, we used the rules from the theoretical section to have new entrants bid at cost, and existing generation bid at 0. In this scenario, it was possible to achieve 0 hours lost, as generators no longer needed to receive extremely high prices from the day-ahead market when they received capacity payments for providing generation resource into the system. Average prices throughout the system were also lower due to increased availability of supply. When factoring in the cost of the capacity payments, the total cost, however, was 40% higher than the optimum found in the energy-only market.

Rather than feeding the optimal bid strategies to the generators as we did earlier, we then let them learn the optimal strategies using a simple reinforcement learner. A few changes were made to the classical agent-based learners seen in the literature. We gave multiple generator agents the same learner, which we called a holding company. This enabled a learner to receive multiple updates per period, since each generator was playng an action and receiving an associated payoff. Because of the nature of our model and the information revealed in the capacity auction we went a step further with the reward function by also updating actions that were not played. This hypothetical reward function coupled with the joint learners allowed us to speed up the convergence of action probabilities.

When the model was run with existing capacity bidding 0 and new entrants learning how to bid, the results were somewhat conclusive. It was very clear that bidding below cost was not beneficial, but there was an even distribution of mass on bids above true cost. The reason for this being that high bids all had an equally unlikely chance of clearing in the auction, thus leading to the same reward payoff. New generation is only built as a result of the capacity auction; no competitive entry occurs. The reserve margin hovers above 1.5% for the duration of the run.

Allowing both new and existing generation to learn the optimal bids provides a rather different distribution over actions of new learners. In this case the distribution peaks at truthful bidding, with the next most likely action being 90% of true costs. This matches up with our findings in Chapter 2, where in the decreasing price auction, the optimal strategy was to bid just shy of the truth. Existing generation had the majority of its probability mass in the 0% to 40% of true costs range. In this scenario, bids and capacity prices in the early years were much more volatile, with some existing generation not clearing in the market from bidding too high. However, as time went on, both the bids and capacity clearing prices stabilized to levels seen in prior runs.

For future analysis there are many modifications that can be made to these scenarios. Different auction structures, such as a discriminatory price auction (where the price paid is equivalent to the bid) could potentially reduce costs to the system but result in wildly different bidding strategies. Here, as we have close to truthful bidding, not much governance or oversight is needed. In a system where results could be manipulated, much more care would need to be taken to ensure actors are playing fair.

The shape of the demand curve used in the capacity auction model is concave, a shape that is rarely seen in classic economics literature. Modifying the shape of the demand curve or adjusting the parameters and observing the effects these have on the market would be an insightful sensitivity analysis.

The original AMES model was used to examine reported marginal supply offers if profit maximizing generators are allowed to bid strategically. They may alter their available capacity or their costs. It was found that generators universally reported higher marginal costs. This version of the AMES model had no new entry or capacity markets. Analysis of what a capacity market does to generator marginal cost reporting would be a valuable addition to the original contributions of the AMES model.

In this paper we found that in a decreasing price style capacity auction results in participants misrepresenting their true costs. We found that when generators behave in this way, capacity markets achieve 100% reliability at a negligible cost to the consumer. Additionally, while good for the consumer, the capacity markets also provided a stable source of income for generators. In the energy only markets, generators were unable to recover their fixed costs, continuously losing money except in the rare hours of the year when prices spiked. Finally, we verified through simulation that generators would learn to bid optimally as we had theorized in the first part of the paper. Thus a complete package is provided for anyone who wishes to do so to utilize the model to run their own analyses, or build upon it as we did.

Appendix A

Generalization of the Single Winner Case

In this section we modify the bounds of the single winner auction to show that the result in the single winner case the bidders undereport their costs is not dependent on a Uniform distribution with the same parameters for both price and costs. We set possible clearing price to be Uniform(0, 1.5) and the costs of the generators to be Uniform(0.75, 1.25). This setup is more similar to the way the current PJM auction is structured. Possible prices in the auction range from 0 to 1.5 times the levelized cost of new entry (CONE). We assume new plants will be mostly centered around this cost of new entry, but with some variation.

For n bidders, we have

$$w'(c) = -\frac{(n-1)(2w(c)-3)(4c-2w(c)-3)}{2(4c-5)(c-w(c))},$$

which again does not have an analytical solution, but does have a numerical solution.

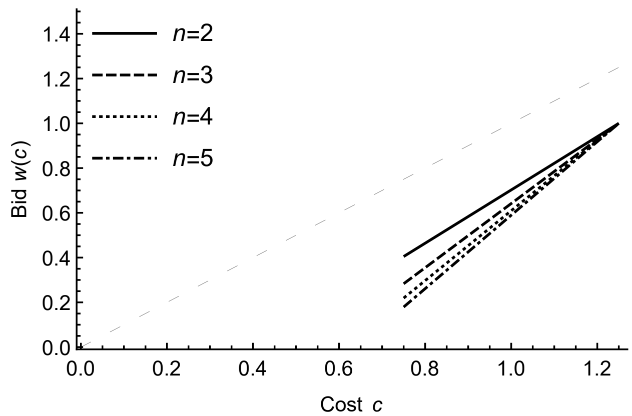


Figure A.1: Optimal bid functions in the single winner case with clearing prices [0, 1.5] and costs [0.75, 1.25].

As n gets large, the bid at the lowest possible cost, $c = 0.75$, goes to 0. The bid at the highest potential cost, $c = 1.25$, is 1, which ensures an expected clearing price of 1.25, since the only way the auction will be won by the bidder will be a price on the interval [1,1.5], and an expected profit of 0.

Generally speaking, there are two cases: one where the maximum possible clearing price is higher than the maximum cost, and one where the maximum possible clearing price is lower than the maximum cost. In the former case, the bid at the maximum cost is equal to the expected clearing price at that bid. The bid is the value $w[\bar{c}]$ that satisfies

$$\bar{c} = \int_{w(\bar{c})}^{b} p h^*(p) \, \mathrm{d}p$$

where $h^*(p)$ is the PDF truncated on the domain $[w(\bar{c}), b]$ of the original clearing price distribution. In the latter case, where the maximum possible clearing price is lower than the maximum possible cost, the optimal bid at the maximum clearing price is $w[b] = b$. For any costs above b, the best strategy is to bid above b, as any possible clearing price would incur a loss. For presentation purposes we will say that $w(c) = c$ when $c > b$. In both cases, there is incentive to under report costs with nearly all distributions.

Bibliography

Ausubel, L. M. and Cramton, P. (2004). Vickrey auctions with reserve pricing. *Economic Theory*, *23*(3), 493–505 (2004).

Blume, A., Heidhues, P., Lafky, J., Mnster, J., and Zhang, M. (2009). All equilibria of the multi-unit vickrey auction. *Games and Economic Behavior*, *66*(2), 729–741. Special Section In Honor of David Gale.

Cramton, P. and Stoft, S. (2007). Colombia firm energy market. *2014 47th Hawaii International Conference on System Sciences*, *0*, 124a.

Engelbrecht-Wiggans, R. and Kahn, C. M. (1998). Multi-unit auctions with uniform prices. *Economic Theory*, *12*(2), 227–258.

Erev, I. and Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *The American Economic Review*, *88*(4), 848–881.

Erev, I. and Roth, A. E. (2007). Multi-agent learning and the descriptive value of simple models. *Artificial Intelligence*, *171*(7), 423 – 428. Foundations of Multi-Agent Learning.

Frayer, J., Keane, S., and Ng, J. (2013). Estimating the value of lost load. Technical report, London Economics Institute.

Gretschko, V., Rasch, A., and Wambach, A. (2014). On the strictly descending multi-unit auction. *Journal of Mathematical Economics*, *50*(0), 79–85.

Hobbs, B., Hu, M.-C., Inon, J., Stoft, S., and Bhavaraju, M. (2007). A dynamic analysis of a demand curve-based capacity market proposal: The pjm reliability pricing model. *Power Systems, IEEE Transactions on*, *22*(1), 3–14.

Hortasu, A. and Puller, S. L. (2008). Understanding strategic bidding in multi-unit auctions: a case study of the texas electricity spot market. *The RAND Journal of Economics*, *39*(1), 86–114.

Joskow, P. L. (2007). Competitive electricity markets and investment in new generating capacity. In D. Helm (Ed.), *The New Energy Paradigm*. Oxford University Press.

Joskow, P. L. (2008). Capacity payments in imperfect electricity markets: Need and design. *Utilities Policy*, *16*(3), 159–170.

Klemperer, P. (2002). What really matters in auction design. *The Journal of Economic Perspectives*, *16*(1), pp. 169–189.

Klemperer, P. (2004). *Auctions: Theory and Practice*. Princeton paperbacks. Princeton University Press.

Krishna, V. (2009). *Auction Theory* (second ed.). Academic Press.

Larsen, E. R., Dyner, I., V, L. B., and Franco, C. J. (2004). Lessons from deregulation in colombia: successes, failures and the way ahead. *Energy Policy*, *32*(15), 1767–1780.

Mann, N., Tsai, C.-H., Glen, G., Schneider, E., Cuevas, P., Dyer, J., Butler, J., Zhang, T., Baldick, R., Deetjen, T., and Morneau, R. (2017). Capacity expansion and dispatch modeling: Model documentation and results for ercot scenarios. Technical report, The Unversity of Texas at Austin Energy Institute.

Martinez-Pardina, I. and Romeu, A. (2011). The case for multi-unit single-run descending-price auctions. *Economics Letters*, *113*(3), 310–313.

Milligan, M. et al. (2011). Methods to model and calculate capacity contributions of variable generation for resource adequacy planning. Technical report, North American Electric Reliability Corporation.

Namovicz, C. (2013). Assessing the economic value of new utility-scale renewable generation projects. EIA Energy Conference.

Ott, A. L. (2014). Re: Vrr curve key parameter review, pjm preliminary recommendations. [Online; accessed 22-October-2015].

PJM (2014). 2017/2018 rpm base residual auction planning period parameters. [Online; accessed 15-October-2015].

Rapoport, A. and Chammah, A. M. (1966). The game of chicken. *American Behavioral Scientist*, *10*(3), 10–28.

Roth, A. E. and Erev, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior*, *8*(1), 164–212.

Schwenen, S. (2015). Strategic bidding in multi-unit auctions with capacity constrained bidders: the new york capacity market. *The RAND Journal of Economics*, *46*(4), 730–750.

Sener, A. C. and Kimball, S. (2007). Reviewing progress in pjm's capacity market structure via the new reliability pricing model. *The Electricity Journal*, *20*(10), 40–53.

Sun, J. and Tesfatsion, L. (2007). Dynamic testing of wholesale power market designs: An open-source agent-based framework. *Computational Economics*, *30*(3), 291–327.

Surendan, R., Hogan, W. W., Hui, H., and Yu, C.-N. (2016). Scarcity pricing in ercot. FERC Technical Conference.

Swinkels, J. M. (2001). Efficiency of large private value auctions. *Econometrica*, *69*(1), 37–68.

Vickrey, W. (1961). Counterspeculation, auctions, and competitive sealed tenders. *The Journal of Finance*, *16*(1), 8–37.